

B14

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number  
**WO 02/086443 A2**

- (51) International Patent Classification<sup>7</sup>: **G01N**
- (21) International Application Number: **PCT/US02/12476**
- (22) International Filing Date: **18 April 2002 (18.04.2002)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
- |            |                               |    |
|------------|-------------------------------|----|
| 60/284,770 | 18 April 2001 (18.04.2001)    | US |
| 60/290,492 | 10 May 2001 (10.05.2001)      | US |
| 60/339,245 | 9 November 2001 (09.11.2001)  | US |
| 60/350,666 | 13 November 2001 (13.11.2001) | US |
| 60/334,370 | 29 November 2001 (29.11.2001) | US |
| 60/372,246 | 12 April 2002 (12.04.2002)    | US |
- (71) Applicant (for all designated States except US): **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **AZIZ, Natasha** [US/US]; 411 California Avenue, Palo Alto, CA 94306 (US). **MURRAY, Richard** [US/US]; 22643 Woodbridge Court, Cupertino, CA 95014 (US).
- (74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— of inventorship (Rule 4.17(iv)) for US only

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

(57) Abstract: Described herein are methods and compositions that can be used for diagnosis and treatment of lung cancer and similar pathologies. Also described herein are methods that can be used to identify modulators of lung cancer and similar pathologies.



WO 02/086443 A2

## METHODS OF DIAGNOSIS OF LUNG CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF LUNG CANCER

5

### CROSS-REFERENCES TO RELATED APPLICATIONS

This application is related to USSN 60/284,770, filed April 18, 2001; USSN 60/290,492, filed May 10, 2001; USSN 60/334,370, filed November 29, 2001; USSN 60/339,245, filed November 9, 2001; USSN 60/350,666, filed November 13, 2001; and  
10 USSN 60/xxx,xxx, filed April 12, 2002 (Docket OMNI-002P); each of which is incorporated herein by reference in its entirety.

### FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression  
15 profiles and nucleic acids, products, and antibodies thereto that are involved in lung cancer; and to the use of such expression profiles and compositions in diagnosis and therapy of lung cancer. The invention further relates to methods for identifying and using agents and/or targets that inhibit lung cancer or related conditions.

20

### BACKGROUND OF THE INVENTION

Lung cancer is the second most commonly occurring cancer in the United States and is the leading cause of cancer-related death. It is estimated that there are over 160,000 new cases of lung cancer in the United States every year. Of those who are diagnosed with lung cancer, 86 percent will die within five years. Lung cancer is the most common visceral  
25 cancer in men and accounts for nearly one third of all cancer deaths in both men and women. In fact, lung cancer accounts for 7% of all deaths, due to any cause, in both men and women.

Smoking is the primary cause of lung cancer, with more than 80% of lung cancers resulting from smoking. About 400 to 500 separate gaseous substances are present in the smoke of a non-filter cigarette. The most noteworthy substances include nitrogen oxides,  
30 hydrogen cyanide, formaldehyde, benzene, and toluene. The particles present in cigarette smoke contain at least 3,500 individual compounds such as nicotine, tobacco alkaloids (nornicotine, anatabine, anabasine), polycyclic aromatic hydrocarbons (e.g., benzo(a)pyrene, B(a)P), naphthalenes, aromatic amines, phenols, and tobacco-specific nitrosamines.



Tobacco-specific nitrosamines are formed during tobacco curing and processing, and are suspected of causing lung cancer in humans. In rodent studies, regardless of the where or how it is applied, the tobacco-specific nitrosamine known as NNK produces lung adenomas and lung adenocarcinomas. The tobacco-specific nitrosamine known as NNAL also produces lung adenocarcinomas in rodents.

Many of the chemicals found in cigarette smoke also affect the nonsmoker inhaling "secondhand" or sidestream smoke. Indeed, the smoke inhaled by non-smokers has a chemical composition similar to the smoke inhaled by smokers, but, importantly, the concentrations of the carcinogenic tobacco-specific nitrosamines are present in higher concentrations in second hand smoke. For this and other reasons, "passive smoking" is an important cause of lung cancer, causing as many as 3,000 lung cancer deaths in nonsmokers each year.

In addition to smoking, other factors thought to be causes of lung cancer include on-the-job exposure to carcinogens such as asbestos and uranium, exposure to chemical hazards such as radon, polycyclic aromatic hydrocarbons, chromium, nickel, and inorganic arsenic, genetic factors, and diet.

Histological classification of various lung cancers define the types of cancer that begin in the lung. See, e.g., Travis, et al. (1999) Histological Typing of Lung and Pleural Tumours (International Histological Classification of Tumours, No 1. Four major cell types make up more than 88% of all primary lung neoplasms. These are: squamous or epidermoid carcinoma, small cell (also called oat cell) carcinoma, adenocarcinoma, and large cell (also called large cell anaplastic) carcinoma. The remainder include undifferentiated carcinomas, carcinoids, bronchial gland tumors, and other rarer types. The various cell types have different natural histories and responses to therapy, and, thus, a correct histologic diagnosis is the first step of effective treatment.

Small cell lung cancer (SCLC) accounts for 18-25% of all lung cancers, and occurs less frequently than non-small cell lung cancers, and generally spread to distant organs more rapidly than non-small cell lung cancer. In general, at the time of presentation small cell lung cancers have already spread beyond the bounds where surgery and curative intent can be undertaken. However, if identified early enough, these cancers are often responsive to chemotherapy and thoracic radiation treatment.

Non-small cell lung cancers (NSCLC) are the more frequently occurring form of lung cancer. They comprise squamous cell carcinoma, adenocarcinoma, and large cell carcinoma

5       The screening of asymptomatic persons at high risk for lung cancer has often proven  
ineffective. In general, only 5 to 15 percent of lung cancer patients have their disease  
detected while they are asymptomatic. Of course, early detection and treatment are critical  
factors in the fight against lung cancer. The average survival rate is 49% for those whose  
cancer is detected early, before the cancer has spread from the lung. Lung cancer often  
0       spreads outside of the lung, and it may have spread to the bones or brain by the time it is  
diagnosed. While the prognosis may be better for lung cancers that are detected early,  
because of the lack of effective curative treatments, early detection does not necessarily alter  
the total death rate from lung cancer.

20

## SUMMARY OF THE INVENTION

25

bronchiectasis. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a lung cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16. Alternatively, the sample may be contacted with a specific binding reagent, e.g., antibody.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16. In another embodiment, the polynucleotide comprises a sequence as shown in Tables 1A-16.

In one embodiment, the biological sample is a tissue sample, or a body fluid. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label. In one embodiment, the polynucleotide is immobilized on a solid surface. In one embodiment, the patient is undergoing a therapeutic regimen to treat lung cancer. In another embodiment, the patient is suspected of having lung cancer. In one embodiment, the patient is a primate, e.g., a human.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy. Or the sample may be evaluated for protein, e.g., contacting the sample with an antibody.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated transcript to a level of the lung cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. Or the sample may be evaluated for comparison of protein.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a

biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated antibody in the biological sample by contacting the biological sample with a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated antibody, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated antibody to a level of the lung cancer-associated antibody in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of lung cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a lung cancer-associated polypeptide in the biological sample by contacting the biological sample with an antibody, wherein the antibody specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby monitoring the efficacy of the therapy.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic treatment. In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1A-16. In one embodiment, an expression vector or cell comprises the isolated nucleic acid. In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-16. In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical. In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

In one aspect, the present invention provides a method of detecting lung cancer in a patient, the method comprising contacting a biological sample from the patient with an antibody or protein as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to a lung cancer gene in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprises a sequence from Tables 1A-16.

5 In another aspect, the present invention provides a method for identifying a compound that modulates a lung cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a lung cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16; and (ii) determining the functional effect of the  
10 compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect. In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant. In one  
15 embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation or another critical process of a lung cancer-associated cell to treat lung cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein. In one embodiment, the compound is an  
20 antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: (i) administering a test compound to a mammal having lung cancer or a cell isolated therefrom; (ii) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables  
25 1A-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of lung cancer.

In one embodiment, the control is a mammal with lung cancer or a cell therefrom that has not been treated with the test compound. In another embodiment, the control is a normal  
30 cell or mammal, or a non-malignant lung disease.

In another aspect, the present invention provides a method for treating a mammal having lung cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having lung cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

## DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and treatment of lung disease or cancer, as well as methods for screening for compositions which modulate lung cancer. "Treatment, monitoring, detection or modulation of lung disease or cancer" includes treatment, monitoring, detection, or modulation of lung disease in those patients who have lung disease (whether malignant or non-malignant, e.g., emphysema, bronchitis, or fibrosis) as well as patients with lung cancers in which gene expression from a gene in Tables 1A-16 is increased or decreased, indicating that the subject is more likely to have disease. In particular, while these targets are identified primarily from lung cancer samples, these same targets are likely to be similarly found in analyses of other medical conditions. These other conditions may result from similar pathological processes which affect similar tissues, e.g., lung cancer, small cell lung carcinoma (oat cell carcinoma), non-small cell carcinomas (e.g., squamous cell carcinoma, adenocarcinoma, large cell lung carcinoma, carcinoid, granulomatous), fibrosis (idiopathic pulmonary fibrosis (IPF), hypersensitivity pneumonitis (HP), interstitial pneumonitis, nonspecific idiopathic pneumonitis (NSIP)), chronic obstructive pulmonary disease (COPD, e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, and esophageal cancer. See, e.g., the NCI webpage and USSN 60/347,349 and USSN 60/xxx,xxx (docket LFBR-001-1P, filed March 29, 2002), each of which is incorporated herein by reference. The treatment may be of lung cancer or related condition itself, or treatment of metastasis.

In particular, identification of markers selectively expressed on these cancers allows for use of that expression in diagnostic, prognostic, or therapeutic methods. As such, the invention defines various compositions, e.g., nucleic acids, polypeptides, antibodies, and small molecule agonists/antagonists, which will be useful to selectively identify those markers. For example, therapeutic methods may take the form of protein therapeutics which use the marker expression for selective localization or modulation of function (for those markers which have a causative disease effect), for vaccines, identification of binding partners, or antagonism, e.g., using antisense or RNAi. The markers may be useful for molecular characterization of subsets of lung diseases, which subsets may actually require

very different treatments. Moreover, the markers may also be important in related diseases to the specific cancers, e.g., which affect similar tissues in non-malignant diseases, or have similar mechanisms of induction/maintenance. Metastatic processes or characteristics may also be targeted. Diagnostic and prognostic uses are made available, e.g., to subset related but distinct diseases, or to determine treatment strategy. The detection methods may be based upon nucleic acid, e.g., PCR or hybridization techniques, or protein, e.g., ELISA, imaging, IHC, etc. The diagnosis may be qualitative or quantitative, and may detect increases or decreases in expression levels.

Tables 1A-16 provide unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in lung cancer samples. The tables also provide an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. In Table 1A, genes marked as "target 1" or "target 2" are particularly useful as therapeutic targets. Genes marked as "target 3" are particularly useful as diagnostic markers. Genes marked as "chron" are upregulated in chronically diseased lung (e.g., emphysema, bronchitis, fibrosis) relative to lung tumors and normal tissue. In certain analyses, the ratio for the "chron" category was determined using the 70th percentile of chronically diseased lung samples divided by the 90th percentile of normal lung samples. The ratio for the targets was determined using the 70th percentile of lung tumor samples divided by the 90th percentile of normal lung samples.

## Definitions

The term "lung cancer protein" or "lung cancer polynucleotide" or "lung cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1A-16 and conservatively modified variants thereof; or (4)

have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "lung cancer polypeptide" and a "lung cancer polynucleotide," include both naturally occurring or recombinant forms.

A "full length" lung cancer protein or nucleic acid refers to a lung cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains the elements normally contained in one or more naturally occurring, wild type lung cancer polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translational processing or splicing, including alternative splicing.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a lung cancer protein, polynucleotide, or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, archival materials, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc.

Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or other mammal; or a bird; reptile; fish. Livestock and domestic animals are of interest.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention in vivo. Archival tissues or materials, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the



same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using, e.g., a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or insertions, substitutions, and naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer

Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology.

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between

two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be negative large numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection catalog or web site, [www.atcc.org](http://www.atcc.org)).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least about 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant or component from the composition to be purified.

In this sense, purification does not require that the purified compound be homogeneous, e.g., 100% pure.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline,  $\gamma$ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refer to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an  $\alpha$  carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain some basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refer to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function similarly to another amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU each encode the amino acid alanine. Thus, at each position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of

conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. In certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally similar molecule. Accordingly, a silent variation of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not necessarily with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. Typically conservative substitutions include for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, Proteins (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (1994) Molecular Biology of the Cell (3<sup>rd</sup> ed.) and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules. "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of  $\beta$ -sheet and  $\alpha$ -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have at least one different linkage, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, in Sanghui and Cook, eds. Carbohydrate Modifications in Antisense Research, ASC Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature ( $T_m$ ) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4° C drop in  $T_m$  for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary

strand; thus the sequences described herein also provide the complement of the sequence.

The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine

5 hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g., the individual units of a peptide nucleic  
10 acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, physiological, chemical, or other physical means. For example, useful labels include  $^{32}\text{P}$ , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins  
15 or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The labels may be incorporated into the cancer nucleic acids, proteins, and antibodies. Many methods known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter, et al. (1962) Nature 144:945; David, et al. (1974) Biochemistry  
20 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth., 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody.

25 The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either  
30 covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method

using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, e.g., through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, preferably one that does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled, e.g., with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled, e.g., with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered



recombinant for the purposes of the invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is typically an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, e.g., wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed in operable linkage to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule selectively to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to essentially no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes (vol. 24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is typically at least two times background, preferably 10 times background hybridization.

Exemplary stringent hybridization conditions are often: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32° C and 48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50° C to about 65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90° C - 95° C for 0.5 - 2 min., an annealing phase lasting 0.5 - 2 min., and an extension phase of about 72° C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al.(1990) PCR Protocols, A Guide to Methods and Applications.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This

occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary “moderately stringent hybridization conditions” include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is at least twice background. Alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., Ausubel, et al. (ed.) Current Protocols in Molecular Biology Lippincott.

The phrase “functional effects” in the context of assays for testing compounds that modulate activity of a lung cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the lung cancer protein or nucleic acid, e.g., a physiological, enzymatic, functional, physical, or chemical effect, such as the ability to decrease lung cancer. It includes ligand binding activity; cell viability, cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of lung cancer cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities.

By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a lung cancer protein sequence, e.g., physiological, functional, enzymatic, physical, or chemical effects. Such functional effects can be measured by many means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the lung cancer protein; measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on lung cancer can also be performed using lung cancer assays known to those of skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein

expression in cells undergoing metastasis, and other characteristics of lung cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for lung cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase,  $\beta$ -gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

"Inhibitors", "activators", and "modulators" of lung cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of lung cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of lung cancer proteins, e.g., antagonists. Antisense or inhibitory nucleic acids may seem to inhibit expression and subsequent function of the protein.

"Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate lung cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of lung cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the lung cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of lung cancer can also be identified by incubating lung cancer cells with the test compound and determining increases or decreases in the expression of 1 or more lung cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more lung cancer proteins, such as lung cancer proteins encoded by the sequences set out in Tables 1A-16.

Samples or assays comprising lung cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a lung cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more

preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., Freshney (1994) Culture of Animal Cells a Manual of Basic Technique pp. 231-241 (3<sup>rd</sup> ed.).

"Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

"Cancer cells," "transformed" cells, or "transformation" in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy (see, Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3<sup>rd</sup> ed.)).

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul, Fundamental Immunology.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible

for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce  $F(ab)'_2$ , a dimer of Fab which itself is a light chain joined to  $V_H$ - $C_H1$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1999) Fundamental Immunology (4th ed.). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; Cole, et al. (1985), pp. 77-96 in Monoclonal Antibodies and Cancer Therapy; Coligan (1991 and supplements) Current Protocols in Immunology; Harlow and Lane (1988) Antibodies. A Laboratory Manual; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783).

A "chimeric antibody" is an antibody molecule in which, e.g., (a) the constant region, or a portion thereof, is altered, replaced, or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function, and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the

variable region, or a portion thereof, is altered, replaced, or exchanged with a variable region having a different or altered antigen specificity.

### Identification of lung cancer-associated sequences

5           In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is  
10       characteristic of the state of the cell. That is, normal tissue may be distinguished from cancerous or metastatic cancerous tissue, or metastatic cancerous tissue can be compared with tissue from surviving cancer patients. By comparing expression profiles of tissue in known different lung cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.  
15       Molecular profiling may distinguish subtypes of a currently collective disease designation, e.g., different forms of lung cancer (chronic disease, adenocarcinoma, etc.)

          The identification of sequences that are differentially expressed in lung cancer versus non-lung cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-  
20       regulate lung cancer, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Malignant disease may be compared to non-malignant conditions. Metastatic tissue can also be analyzed to determine  
25       the stage of lung cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the lung cancer expression profile. This may be done by making biochips comprising sets of the important lung cancer genes,  
30       which can then be used in these screens. PCR methods may be applied with selected primer pairs, and analysis may be of RNA or of genomic sequences. These methods can also be done on the protein basis; that is, protein expression levels of the lung cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the lung cancer

nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the lung cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs or as protein or DNA vaccines.

5        Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in lung cancer relative to normal tissues and/or non-malignant lung disease, or in different types of lung disease, herein termed "lung cancer sequences." As outlined below, lung cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in lung cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the lung cancer sequences are from humans; however, as will be appreciated by those in the art, lung cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other lung cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets (dogs, cats, etc.). Lung cancer sequences from other organisms may be obtained using the techniques outlined below.

10        Lung cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, lung cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the lung cancer sequences can be generated.

20        A lung cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, e.g., using homology programs or hybridization conditions.

30        For identifying lung cancer-associated sequences, the lung cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, cancer and non-malignant conditions, non-malignant conditions and normal tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing lung cancer samples with metastatic cancer samples from other cancers, such as, breast, other gastrointestinal cancers, prostate, ovarian,



etc. Samples of, non metastatic disease tissue and tissue undergoing metastasis are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g., from Affymetrix, Santa Clara, CA. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal lung, but also including, and not limited to colon, heart, brain, liver, breast, kidney, muscle, prostate, small intestine, large intestine, spleen, bone, and/or placenta. In a preferred embodiment, those genes identified during the lung cancer screen that are expressed in significant amounts in other tissues (e.g., essential organs) are removed from the profile, although in some embodiments, this is not necessary (e.g., where organs may be dispensable at a later stage of life). That is, when screening for drugs, it is usually preferable that the target expression be disease specific, to minimize possible side effects on other organs.

In a preferred embodiment, lung cancer sequences are those that are up-regulated in lung cancer; that is, the expression of these genes is higher in cancerous tissue than in normal lung or other tissue. "Up-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, DA, et al (1998) *Nucleic Acids Research* 26:1-7 and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. In some situations, the sequences may be derived from assembly of available sequences or be predicted from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev (2000) *Genome Res.* 10:516-522). In other situations, sequences have been derived from cloning and sequencing of isolated nucleic acids.

In another preferred embodiment, lung cancer sequences are those that are down-regulated in the lung cancer; that is, the expression of these genes is lower in cancerous tissue

or normal lung or other tissue. "Down-regulation" as used herein means, when the ratio is presented as a number greater than one, that the ratio is greater than one, preferably 1.5 or greater, more preferably 2.0 or greater, or, when the ratio is presented as a number less than one, that the ratio is less than one, preferably 0.5 or less, more preferably 0.25 or less.

### Informatics

The ability to identify genes that are over or under expressed in lung cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with lung cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson (1998) Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample representing lung cancer, i.e., the identification of lung cancer-associated sequences described herein, provide an abundance of information, which can be correlated with

pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount, et al. (2001) Bioinformatics; Durbin, et al. (eds., 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids (; Baxeavanis and Oeullette (eds., 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins); Rashidi and Buehler (1999) Bioinformatics: Basic Applications in Biological Science and Medicine; Setubal, et al. (eds 1997) Introduction to Computational Molecular Biology; Misener and Krawetz (eds, 2000) Bioinformatics: Methods and Protocols; Higgins and Taylor (eds., 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical

Approach; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet; Han and Kamber (2000) Data Mining: Concepts and Techniques (2000); and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes.

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for lung cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may

be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux,  
5 SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices  
10 linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

15 The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

20 In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results.

25 Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary  
memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or  
30 SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC,

MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

#### **Characteristics of lung cancer-associated proteins**

Lung cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the lung cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Alberts (ed. 1994) Molecular Biology of the Cell (3d ed.)). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2

domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by

5 one of ordinary skill in the art, these motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains.

10 Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman, et al (2000) *Nuc. Acids Res.* 28:263-266; Sonnhammer, et al. (1997) *Proteins* 28:405-420; Bateman, et al. (1999) *Nuc. Acids Res.* 27:260-262; and Sonnhammer, et al. (1998) *Nuc. Acids Res.* 26:320-322).

15 In another embodiment, the lung cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

25 Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels, pumps, and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the

localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site <http://psort.nibb.ac.jp/>).

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, hormones, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they may mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains may also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Lung cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for extracellular immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ* or in histological analysis. Alternatively, antibodies can also label intracellular proteins, in which case analytical samples are typically permeabilized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful lung markers of disease.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the lung cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; e.g., if circulating, they often serve to transmit



signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g., secretion into the blood stream), or exocrine (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, sex producing glands of the ear, etc.). Thus secreted molecules often find use in modulating or altering numerous aspects of physiology. Lung cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

#### Use of lung cancer nucleic acids

As described above, lung cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the lung cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The lung cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1A-16, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the lung cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, et al., *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once a lung cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire lung cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant lung cancer nucleic acid can be further-used as a probe to identify and isolate

other lung cancer nucleic acids, e.g., extended coding regions. It can also be used as a “precursor” nucleic acid to make modified or variant lung cancer nucleic acids and proteins.

The lung cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the lung cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, RNAi, vaccine, and/or antisense applications. Alternatively, the lung cancer nucleic acids that include coding regions of lung cancer proteins can be put into expression vectors for the expression of lung cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to lung cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the lung cancer nucleic acids, i.e., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by “substantially complementary” herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under appropriate reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally complements of ORFs or whole genes are not used. In some embodiments, nucleic acids of lengths up to hundreds of bases can be used.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a

particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is typically meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to a biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. Often the substrate may contain discrete individual sites appropriate for individual partitioning and identification. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in US application entitled Reusable Low Fluorescent Plastic Biochip, U.S.

Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers as are known in the art; e.g., homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized, and then attached to the surface of the solid support. Either the 5' or 3' terminus may be attached to the solid support, or attachment may be via linkage to an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized *in situ*, using known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of lung cancer-associated sequences. These assays are typically performed in conjunction with

reverse transcription. In such assays, a lung cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of lung cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols, A Guide to Methods and Applications.

In some embodiments, a TaqMan based assay is used to measure expression.

TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer, e.g., [www2.perkin-elmer.com](http://www2.perkin-elmer.com)).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560, Landegren, et al. (1988) Science 241:1077, and Barringer, et al. (1990) Gene 89:117), transcription amplification (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173), self-sustained sequence replication (Guatelli, et al. (1990) Proc. Nat. Acad. Sci. USA 87:1874), dot PCR, and linker adapter PCR, etc.

#### **Expression of lung cancer proteins from nucleic acids**

In a preferred embodiment, lung cancer nucleic acids, e.g., encoding lung cancer proteins, are used to make a variety of expression vectors to express lung cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (see, e.g., Ausubel, *supra*, and Fernandez and Hoeffler (eds 1999) Gene Expression Systems) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the lung cancer protein. The term "control sequences" refers to DNA

sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding  
10 sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction  
15 sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the lung cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are  
20 not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences may be either constitutive or inducible promoters. The promoters  
25 may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two  
30 organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector often contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating

vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez and Hoeffler, *supra*).

In addition, in a preferred embodiment, the expression vector contains a selectable  
5 marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The lung cancer proteins of the present invention are usually produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a lung cancer protein, under the appropriate conditions to induce or cause expression of the lung  
10 cancer protein. Conditions appropriate for lung cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate  
15 growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae*  
20 and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the lung cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and  
25 adenoviral systems. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez and Hoeffler, *supra*). Typically, transcription termination and  
30 polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, lung cancer proteins are expressed in bacterial systems. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the lung cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez and Hoeffer, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, lung cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, lung cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*,



*Kluyveromyces fragilis* and *K. lactis*, *Pichia guillermondii*, and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The lung cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the lung cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the lung cancer protein may be made as a fusion protein to increase expression for affinity purification purposes, or for other reasons. For example, when the lung cancer protein is a lung cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the lung cancer protein is purified or isolated after expression. Lung cancer proteins may be isolated or purified in a variety of appropriate ways. Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the lung cancer protein may be purified using a standard anti-lung cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes (1982) Protein Purification. The degree of purification necessary will vary depending on the use of the lung cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the lung cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, therapeutic entities, for production of antibodies, as transcription or translation inhibitors, etc.

## 25 Variants of lung cancer proteins

In one embodiment, the lung cancer proteins are derivative or variant lung cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative lung cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at a particular residue within the lung cancer peptide.

Also included within one embodiment of lung cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three

classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the lung cancer protein, using cassette or PCR mutagenesis or other techniques, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above.

5 However, variant lung cancer protein fragments having up to about 100-150 residues may be prepared by *in vitro* synthesis. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the lung cancer protein amino acid sequence. The variants typically exhibit a similar qualitative biological activity as the naturally occurring analogue,

10 although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is often predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be

15 conducted at the target codon or region and the expressed lung cancer variants screened for the optimal combination of desired activity. Techniques exist for making substitution mutations at predetermined sites in DNA having a known sequence, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of mutants is often done using assays of lung cancer protein activities.

20 Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be occasionally tolerated. Deletions generally range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive

25 at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. Larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of a lung cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution chart provided in the definition section.

30 Variants typically exhibit essentially the same qualitative biological activity and will elicit the same immune response as a naturally-occurring analog, although variants also are selected to modify the characteristics of lung cancer proteins as needed. Alternatively, the

variant may be designed or reorganized such that a biological activity of the lung cancer protein is altered. For example, glycosylation sites may be added, altered, or removed.

Covalent modifications of lung cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid

5 residues of a lung cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a lung cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking lung cancer polypeptides to a water-insoluble support matrix or surface for use in a method for purifying anti-lung cancer polypeptide antibodies or screening assays, as is more fully  
10 described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimideate.

15 Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl or tyrosyl residues, methylation of the  $\gamma$ -amino groups of lysine, arginine, and histidine side chains (Creighton (1983) Proteins: Structure and Molecular Properties, pp. 79-86), acetylation of the N-terminal  
20 amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the lung cancer polypeptide encompassed by this invention is an altered native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended herein to mean adding to or deleting one or more carbohydrate moieties of a native sequence lung cancer polypeptide. Glycosylation patterns  
25 can be altered in many ways. For example the use of different cell types to express lung cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to lung cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence  
30 lung cancer polypeptide (for O-linked glycosylation sites). The lung cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the lung cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the lung cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin and Wriston (1981) CRC Crit. Rev. Biochem., pp. 259-306.

5 Removal of carbohydrate moieties present on the lung cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al. (1987) Arch. Biochem. Biophys., 259:52 and by Edge, et al. (1981) Anal. Biochem., 118:131.

10 Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura, et al. (1987) Meth. Enzymol., 138:350.

Another type of covalent modification of lung cancer comprises linking the lung cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent  
15 Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192, or 4,179,337.

Lung cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a lung cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric  
20 molecule comprises a fusion of a lung cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the lung cancer polypeptide. The presence of such epitope-tagged forms of a lung cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the lung  
25 cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a lung cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

30 Various tag polypeptides and their respective antibodies are well known and examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies

thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3(6):547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:6393-6397).

Also included are other lung cancer proteins of the lung cancer family, and lung cancer proteins from other organisms, which are cloned and expressed as outlined below.

Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related lung cancer proteins from primates or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include unique areas of the lung cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. PCR reaction conditions are well known in the art (e.g., Innis, PCR Protocols, *supra*).

#### Antibodies to lung cancer proteins

In a preferred embodiment, when a lung cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the lung cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller lung cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are well known (e.g., Coligan, *supra*; and Harlow and Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of Tables 1A-16 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal

being immunized. Immunogenic proteins include, e.g., keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Adjuvants include, e.g., Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) Nature 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of the tables, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding (1986) Monoclonal Antibodies: Principles and Practice, pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovin, or primate origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are typically monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of the tables or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to lung cancer protein are capable of reducing or eliminating a biological function of a lung cancer protein, in a naked form or conjugated to an effector moiety. That is, the addition of anti-lung cancer protein antibodies (either polyclonal or preferably monoclonal) to lung cancer tissue (or cells containing lung cancer) may reduce or eliminate the lung cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the lung cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Medarex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of a human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. A humanized antibody optimally also will typically comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-329; and Presta (1992) Curr. Op. Struct. Biol. 2:593-596). Humanization can be performed following the method of Winter and co-workers (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-327; Verhoeyen, et al. (1988) Science 239:1534-1536), by substituting rodent CDRs or CDR sequences for corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by corresponding sequence from a non-human species.

Human-like antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom and Winter (1991) J. Mol. Biol. 227:381; Marks, et al. (1991) J. Mol. Biol. 222:581). The techniques of Cole, et al. and Boerner, et al. are also available for the preparation of human monoclonal antibodies (Cole, et al. (1985)

- 5 Monoclonal Antibodies and Cancer Therapy, p. 77 and Boerner, et al. (1991) J. Immunol. 147(1):86-95). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in
- 10 nearly all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks, et al. (1992) Bio/Technology 10:779-783; Lonberg, et al. (1994) Nature 368:856-859; Morrison (1994) Nature 368:812-13; Fishwild, et al. (1996) Nature Biotechnology 14:845-51; Neuberger
- 15 (1996) Nature Biotechnology 14:826; and Lonberg and Huszar (1995) Intern. Rev. Immunol. 13:65-93.

By immunotherapy is meant treatment of lung cancer with an antibody raised against a lung cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient).

- 20 Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. The antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for
- 25 expression of the antigen, leading to an immune response.

In a preferred embodiment the lung cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, may bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted lung cancer protein.

- 30 In another preferred embodiment, the lung cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment may bind the extracellular domain of the lung cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The



antibody may cause down-regulation of the transmembrane lung cancer protein. The antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the lung cancer protein. The antibody may be an antagonist of the lung cancer protein or may prevent activation of a transmembrane lung cancer protein, or may induce or suppress a particular cellular pathway. In some embodiments, when the antibody prevents the binding of other molecules to the lung cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF- $\alpha$ , TNF- $\beta$ , IL-1, INF- $\gamma$ , and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody may belong to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, lung cancer may be treated by administering to a patient antibodies directed against the transmembrane lung cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be various molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of a lung cancer protein. In another aspect the therapeutic moiety may modulate an activity of molecules associated with or in close proximity to a lung cancer protein. The therapeutic moiety may inhibit enzymatic or signaling activity such as protease or collagenase activity associated with lung cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to lung cancer tissue or cells results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with lung cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, croton, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against lung cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane lung cancer proteins not only serves to increase the local concentration of therapeutic moiety in the lung cancer

afflicted area, but also serves to reduce deleterious side effects that may be associated with the untargeted therapeutic moiety.

In another preferred embodiment, the lung cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein or other entity which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the lung cancer protein can be targeted within a cell, i.e., the nucleus, an antibody theretomay contain a signal for that target localization, i.e., a nuclear localization signal.

The lung cancer antibodies of the invention specifically bind to lung cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a  $K_d$  of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most preferably, 0.01  $\mu$ M or better. Selectivity of binding to the specific target and not to related other sequences is also important.

#### Detection of lung cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the lung cancer phenotype. Expression levels of genes in normal tissue (e.g., not undergoing lung cancer), in lung cancer tissue (and in some cases, for varying severities of lung cancer that relate to prognosis, as outlined below), or in non-malignant disease are evaluated to provide expression profiles. A gene expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state of the cell. While two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g.,

normal versus lung cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart (1996) Nature Biotechnology 14:1675-1680, hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is typically at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the RNA or DNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the lung cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to lung cancer genes, e.g., those identified as being important in a lung cancer or disease phenotype, can be evaluated in a lung cancer diagnostic test. In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes.

The lung cancer nucleic acid probes may be attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In a preferred embodiment nucleic acids encoding the lung cancer protein are detected. Although DNA or RNA encoding the lung cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a lung cancer protein is detected.

Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed *in situ*. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a lung cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, lung cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as markers of lung cancer, e.g., for prognostic or diagnostic purposes. Detection of these proteins in putative lung cancer tissue allows for detection, prognosis, or diagnosis of lung cancer or similar disease, and perhaps for selection of therapeutic strategy. In one embodiment, antibodies are used to detect lung cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the lung cancer protein is detected, e.g., by immunoblotting with antibodies raised against the lung cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the lung cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., Asai (ed. 1993) Methods in Cell Biology:

Antibodies in Cell Biology, volume 37. In this method cells are contacted with from one to many antibodies to the lung cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label, e.g., multicolor fluorescence or confocal imaging. In another method the primary antibody to the lung cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of lung cancer proteins. Many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing lung cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of lung cancer proteins. Antibodies can be used to detect a lung cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous lung cancer protein or vaccine.

In a preferred embodiment, *in situ* hybridization of labeled lung cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including lung cancer tissue and/or normal tissue, are made. *In situ* hybridization (see, e.g., Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to lung cancer, clinical, pathological, or other information, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple

genes may be useful in various combinations. As above, lung cancer probes may be attached to biochips for the detection and quantification of lung cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

5

### Assays for therapeutic compounds

In a preferred embodiment, the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing lung cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) Science 279:84-8; Heid (1996) Genome Res. 6:986-94.

10

15

In a preferred embodiment, the lung cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified lung cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the lung cancer phenotype or an identified physiological function of a lung cancer protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

20

25

Having identified differentially expressed genes herein, a variety of assays may be performed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene with altered regulation in lung cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the lung cancer protein. "Modulation" thus includes an increase or a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing lung cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in lung cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in lung

30

cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the lung cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene or protein expression monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the lung cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of lung cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more lung cancer-associated sequences, e.g., a polynucleotide sequence set out in the tables. Generally, in a preferred embodiment, a test compound is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate lung cancer, modulate lung cancer proteins, bind to a lung cancer protein, or interfere with the binding of a lung cancer protein and an antibody, substrate, or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the lung cancer phenotype or the expression of a lung cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles of nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a lung cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induces a lung cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

In one aspect, a modulator will neutralize the effect of a lung cancer protein. By “neutralize” is meant that activity of a protein and the consequent effect on the cell is inhibited or blocked.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a lung cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such “combinatorial chemical libraries” are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop, et al. (1994) J. Med. Chem. 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature, 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Nat. Acad. Sci. USA 90:6909-6913), vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et



al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661), oligocarbamates (Cho, et al. (1993) Science 261:1303), and/or peptidyl phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385, nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14(3):309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. The above devices, with appropriate modification, are suitable for use with the present invention. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, RU, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect modulation of lung cancer gene transcription, polypeptide expression, and polypeptide activity.

High throughput assays for evaluating the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins,

U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g.,

5 Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a  
10 high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or  
15 fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly  
20 useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring  
25 proteins, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that the nucleic acid or peptide consists of essentially random sequences of nucleotides and amino acids, respectively. Since these random peptides (or nucleic acids, discussed below) are often chemically synthesized, they may incorporate a nucleotide or amino acid at any position. The synthetic process can be designed to generate  
30 randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. In a preferred embodiment, the nucleotides or amino acid residues are  
5 randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc.

Modulators of lung cancer can also be nucleic acids, as defined above.

10 As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. Digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

15 After a candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence is analyzed. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription  
20 with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as,  
25 alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin  
30 is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

Nucleic acid assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos.

5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc., which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the lung cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state,

screens are performed to identify agents that bind and/or modulate the biological activity of the gene product, or evaluate genetic polymorphisms.

Genes can be screened for those that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a lung cancer expression pattern leading to a normal expression pattern, or to modulate a single lung cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated lung cancer tissue reveals genes that are not expressed in normal tissue or lung cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for lung cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated lung cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of lung cancer cells, that have an associated lung cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once a test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., lung cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress a lung cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on lung cancer activity. By defining such a signature for the lung cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

Measure of lung cancer polypeptide activity, or of lung cancer or the lung cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the metastatic polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of lung cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian lung cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a lung cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the lung cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is typically measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the lung cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is typically detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a lung cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or  $\beta$ -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed

genes are sometimes referred to herein as "lung cancer proteins." The lung cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another  
5 embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated  
10 gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the lung cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a lung cancer  
15 protein and a candidate compound, and determining the binding of the compound to the lung cancer protein. Preferred embodiments utilize the human lung cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative lung cancer proteins may be used.

20 Generally, in a preferred embodiment of the methods herein, the lung cancer protein or the candidate agent is non-diffusably bound to an insoluble support, preferably having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of  
25 screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents  
30 and samples. The particular manner of binding of the composition is typically not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition, and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation

sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the lung cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the lung cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the lung cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the lung cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g.,  $^{125}\text{I}$  for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor may be a binding moiety known to bind to the target molecule (i.e., a lung cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically



between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by a test compound. Displacement of the competitor is an indication that the test compound is binding to the lung cancer protein and thus is capable of binding to, and potentially modulating, the activity of the lung cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the lung cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the lung cancer protein.

In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the lung cancer proteins. In one embodiment, the methods comprise combining a lung cancer protein and a competitor in a first sample. A second sample comprises a test compound, a lung cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the lung cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the lung cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native lung cancer protein, but cannot bind to modified lung cancer proteins. The structure of the lung cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a lung cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results.

Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a lung cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising lung cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a lung cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate lung cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the lung cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting lung cancer cell division is provided. The method comprises administration of a lung cancer inhibitor. In another embodiment, a method of inhibiting lung cancer is provided. The method may comprise administration of a lung cancer inhibitor. In a further embodiment, methods of treating cells or individuals with lung cancer are provided, e.g., comprising administration of a lung cancer inhibitor.

In one embodiment, a lung cancer inhibitor is an antibody as discussed above. In another embodiment, the lung cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, viability, and metastasis assays are known to those of skill in the art, as described below.

*Soft agar growth or colony formation in suspension*

5 Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and  
10 grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of lung cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

15 Techniques for soft agar growth or colony formation in suspension assays are described in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (3<sup>rd</sup> ed.), herein incorporated by reference. See also, the methods section of Garkavtsev, et al. (1996), *supra*, herein incorporated by reference.

20 *Contact inhibition and density limitation of growth*

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a  
25 higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (<sup>3</sup>H)-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a  
30 normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with (<sup>3</sup>H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a lung cancer-associated sequence and are grown for 24 hours at saturation density in

non-limiting medium conditions. The percentage of cells labeling with ( $^3\text{H}$ )-thymidine is determined autoradiographically. See, Freshney (1994), *supra*.

#### *Growth factor or serum dependence*

5 Transformed cells typically have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Inst. 37:167-175; Eagle, et al. (1970) J. Exp. Med. 131:836-879); Freshney, *supra*. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

#### *Tumor specific markers levels*

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g.,  
15 Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) "Angiogenesis and Cancer" in Sem Cancer Biol.).

Various techniques which measure the release of these factors are described in  
20 Freshney (1994), *supra*. Also, see, Unkeless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" in Mihich (ed. 1985) Biological Responses in Cancer, pp. 178-184; Freshney Anticancer Res. 5:111-130 (1985).

#### *Invasiveness into Matrigel*

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate lung cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and  
30 invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with  $^{125}\text{I}$  and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), *supra*.

#### *Tumor growth in vivo*

Effects of lung cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the lung cancer gene is disrupted or in which a lung cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous lung cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous lung cancer gene with a mutated version of the lung cancer gene, or by mutating the endogenous lung cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) Science 244:1288). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella, et al. (1974) J. Natl. Cancer Inst. 52:921), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) Br. J. Cancer 38:263; Selby, et al. (1980) Br. J. Cancer 41:52) can be used as a host. Transplantable tumor cells (typically about  $10^6$  cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a lung cancer-associated sequences are injected subcutaneously. After a suitable length of time,

preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

## 5 Polynucleotide modulators of lung cancer

### *Antisense and RNAi Polynucleotides*

In certain embodiments, the activity of a lung cancer-associated protein is downregulated, or entirely inhibited, by the use of antisense or an inhibitory polynucleotide, i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a  
10 coding mRNA nucleic acid sequence, e.g., a lung cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their  
15 close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the lung cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

20 Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides.  
25 Sense oligonucleotides can, e.g., be employed to block transcription by binding to the antisense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for lung cancer molecules. A preferred antisense molecule is for a lung cancer sequence in the tables, or for a ligand or activator thereof. Antisense or sense  
30 oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein

is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659 and van der Krol, et al. (1988) BioTechniques 6:958).

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Scienceexpress (21March2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

### *Ribozymes*

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of lung cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (see, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. (1994) Human Gene Therapy 1:39-45; Leavitt, et al. (1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (19994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205: 121-126).

Polynucleotide modulators of lung cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of lung cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by

formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating lung cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-lung cancer antibody that reduces or eliminates the biological activity of an endogenous lung cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a lung cancer protein. This may be accomplished in any number of ways. In a preferred embodiment, e.g., when the lung cancer sequence is down-regulated in lung cancer, such state may be reversed by increasing the amount of lung cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous lung cancer gene or administering a gene encoding the lung cancer sequence, using known gene-therapy techniques. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, e.g., when the lung cancer sequence is up-regulated in lung cancer, the activity of the endogenous lung cancer gene is decreased, e.g., by the administration of a lung cancer antisense or RNAi nucleic acid.

In one embodiment, the lung cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to lung cancer proteins. Similarly, the lung cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify lung cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a lung cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The lung cancer antibodies may be coupled to standard affinity chromatography columns and used to purify lung cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the lung cancer protein.

### **Methods of identifying variant lung cancer-associated sequences**

Without being bound by theory, expression of various lung cancer sequences is correlated with lung cancer. Accordingly, disorders based on mutant or variant lung cancer genes may be determined. In one embodiment, the invention provides methods for



identifying cells containing variant lung cancer genes, e.g., determining all or part of the sequence of at least one endogenous lung cancer genes in a cell. In a preferred embodiment, the invention provides methods of identifying the lung cancer genotype of an individual, e.g., determining all or part of the sequence of at least one lung cancer gene of the individual.

- 5 This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced lung cancer gene to a known lung cancer gene, i.e., a wild-type gene.

10 The sequence of all or part of the lung cancer gene can then be compared to the sequence of a known lung cancer gene to determine if any differences exist. This can be done using known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the lung cancer gene of the patient and the known lung cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

- 15 In a preferred embodiment, the lung cancer genes are used as probes to determine the number of copies of the lung cancer gene in the genome.

In another preferred embodiment, the lung cancer genes are used as probes to determine the chromosomal localization of the lung cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when  
20 chromosomal abnormalities such as translocations, and the like are identified in the lung cancer gene locus.

#### **Administration of pharmaceutical and vaccine compositions**

- In one embodiment, a therapeutically effective dose of a lung cancer protein or  
25 modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel, et al. (1992) Pharmaceutical Dosage Forms and Drug Delivery; Lieberman, Pharmaceutical Dosage Forms (vols. 1-3), Dekker, ISBN 0824770846,  
30 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding; and Pickar (1999) Dosage Calculations). Adjustments for lung cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration,

drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art.

A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and  
5 veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the lung cancer proteins and modulators thereof of the present invention can be done in a variety of ways, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly,  
10 intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the lung cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a lung cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the  
15 pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid,  
20 sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases  
25 such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic  
30 ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose,

lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that lung cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a lung cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., Remington's Pharmaceutical Science (15th ed., 1980) and Hardman, et al. (eds. 1996) Goodman and Gilman: The Pharmacological Basis of Therapeutics).

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, e.g., Remington's Pharmaceutical Science and Goodman and Gilman, The Pharmacological Basis of Therapeutics, *supra*.

The compositions containing modulators of lung cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer based, at least in part, upon gene expression profiles. Vaccine strategies may be used, in either a DNA vaccine form, or protein vaccine.

It will be appreciated that the present lung cancer protein-modulating compounds can be administered alone or in combination with additional lung cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in the tables, such as antisense or RNAi polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of lung cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo*, or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see, e.g.,

Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 (Berger), Ausubel, et al. (eds. 1999) Current Protocols (supplemented through 1999), and Sambrook, et al. (1989) Molecular Cloning - A Laboratory Manual (2nd ed., Vol. 1-3).

5 In a preferred embodiment, lung cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, lung cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the lung cancer coding regions) can be administered in a gene therapy application. These lung cancer genes can include antisense or inhibitory applications, e.g., as inhibitory RNA or gene  
10 therapy (e.g., for incorporation into the genome) or as antisense compositions.

Lung cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses.. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g., Vitiello, et al. (1995) J. Clin. Invest. 95:341), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres  
15 (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. U.S.A. 85:5409-5413; Tam  
20 (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, et al., p. 379 In: Kaufmann (ed. 1996) Concepts in vaccine development; Chakrabarti, et al. (1986) Nature 320:535; Hu, et al. (1986) Nature 320:537; Kieny, et al. (1986) AIDS Bio/Technology 4:790; Top, et al. (1971) J. Infect. Dis. 124:148; Chanda, et al.  
25 (1990) Virology 175:535), particles of viral or synthetic origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25; Eldridge, et al. (1993) Sem. Hematol. 30:16; Falo, et al. (1995) Nature Med. 7:649), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369; Gupta, et al. (1993) Vaccine 11:293), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585; Rock (1996) Immunol. Today 17:131), or, naked or particle absorbed cDNA (Ulmer, et al. (1993)  
30 Science 259:1745; Robinson, et al. (1993) Vaccine 11:957; Shiver, et al., p. 423 In: Kaufmann (ed. 1996) Concepts in vaccine development; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923 and Eldridge, et al. (1993) Sem. Hematol. 30:16). Toxin-targeted

delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff, et. al. (1990) *Science* 247:1465 as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode lung cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) *Nature* 351:456-460. A wide variety of other vectors useful for therapeutic administration or immunization e.g., adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the

like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata, et al. (2000) Mol Med Today 6:66-71; Shedlock, et al. (2000) J. Leukoc. Biol. 68:793-806; Hipp, et al. (2000) In Vivo 14:571-85).

Methods for the use of genes as DNA vaccines are well known, and include placing a lung cancer gene or portion of a lung cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a lung cancer patient. The lung cancer gene used for DNA vaccines can encode full-length lung cancer proteins, but more preferably encodes portions of the lung cancer proteins including peptides derived from the lung cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a lung cancer gene. For example, lung cancer-associated genes or sequence encoding subfragments of a lung cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the lung cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment lung cancer genes find use in generating animal models of lung cancer. When the lung cancer gene identified is repressed or diminished in metastatic tissue, gene therapy technology, e.g., wherein antisense or inhibitory RNA directed to the lung cancer gene will also diminish or repress expression of the gene. Animal models of lung cancer find use in screening for modulators of a lung cancer-associated sequence or modulators of lung cancer. Similarly, transgenic animal technology including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the lung cancer protein. When desired, tissue-specific expression or knockout of the lung cancer protein may be necessary.

It is also possible that the lung cancer protein is overexpressed in lung cancer. As such, transgenic animals can be generated that overexpress the lung cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene.

Animals generated by such methods will find use as animal models of lung cancer and are additionally useful in screening for modulators to treat lung cancer.

#### **Kits for Use in Diagnostic and/or Prognostic Applications**

5           For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In diagnostic and research applications such kits may include at least one of the following: assay reagents, buffers, lung cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, RNAi, dominant negative lung cancer polypeptides or polynucleotides, small molecule inhibitors of  
10 lung cancer-associated sequences, etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

          In addition, the kits may include instructional materials containing instructions (e.g., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. A medium  
15 capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

          The present invention also provides for kits for screening for modulators of lung  
20 cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a lung cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing lung cancer-associated activity. Optionally, the kit contains biologically active lung cancer protein. A wide variety of kits and components can be prepared according to the present  
25 invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes typically will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.



## EXAMPLES

## Example 1: Gene Chip Analysis

Molecular profiles of various normal and cancerous tissues were determined and  
5 analyzed using gene chips. RNA was isolated and gene chip analysis was performed as  
described (Glynne, et al. (2000) Nature 403:672-676; Zhao, et al. (2000) Genes Dev. 14:981-  
993).

Tables 1A and 1B were previously filed on April 18, 2001 in USSN 60/284,770 (18501-001500US) and on November 29, 2001 in USSN 60/334,370 (18501-001520US)

5	Table 1A					
	Pkey	ExAccn	UnigeneID	Unigene Title	70% chron/90% NL	70% SQAD/90% NL
	100134	D13264	Hs.49	macrophage scavenger receptor 1	1.61	0.74
	100780	HG3731-HT4001		***Immunoglobulin Heavy Chain, VdJc Reg	2.68	3.28
10	100971	J02874	Hs.83213	fatty acid binding protein 4; adipocyte	1.96	0.14
	101088	L05568	Hs.553	solute carrier family 6 (neurotransmitter)	0.79	0.07
	101102	L07594	Hs.79059	transforming growth factor; beta receptor	2.55	1
	101168	L15388	Hs.211569	G protein-coupled receptor kinase 5	0.88	0.27
	101277	L38486	Hs.118223	microfibrillar-associated protein 4	0.89	0.26
15	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.59	0.29
	101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene h	1.15	0.41
	101345	L76380	Hs.152175	calctonin receptor-like	0.81	0.31
	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a I	1.31	0.77
	101764	M80583	Hs.81256	S100 calcium-binding protein A4 (calcium	1.44	0.82
20	101771	M81750	Hs.153837	myeloid cell nuclear differentiation ant	0.96	0.45
	101842	M93221	Hs.75182	mannose receptor; C type 1	1.27	0.37
	102283	U31384	Hs.83381	guanine nucleotide binding protein 11	1.04	0.3
	102363	U39447	Hs.198241	amine oxidase; copper containing 3 (vasc	0.96	0.26
	102507	U52154	Hs.193044	potassium inwardly-rectifying channel; s	2.81	3.45
25	102698	U75272	Hs.1867	progastricsin (pepsinogen C)	0.95	0.23
	103025	X54131	Hs.123641	protein tyrosine phosphatase; receptor t	1.62	0.21
	103280	X79881	Hs.76206	cadherin 5; VE-cadherin (vascular epithe	0.9	0.41
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2	1.27	0.49
	103541	Z11697	Hs.79197	CD83 antigen (activated B lymphocytes; I	1.86	1
30	103554	Z18951	Hs.74034	caveolin 1; caveolae protein; 22kD	1.27	0.47
	104212	AB002238	Hs.173035	KIAA0300 protein	1.17	0.16
	104691	AA011176	Hs.37744	ESTs	1.08	0.35
	104825	AA035613	Hs.141883	ESTs	0.75	0.27
	104857	AA043219	Hs.19058	ESTs	2.6	3.3
35	104865	AA045136	Hs.22575	ESTs	1.23	0.49
	104989	AA102038	Hs.118615	ESTs	0.63	0.32
	105729	AA292694	Hs.3807	ESTs; Weakly similar to PHOSPHOLEMMAN PR	0.86	0.34
	105847	AA398606	Hs.32241	ESTs	1.32	0.4
	105894	AA400979	Hs.25691	calctonin receptor-like receptor activi	0.78	0.28
40	106490	AA451861	Hs.115537	ESTs; Weakly similar to dipeptidase prec	1.2	0.47
	106536	AA453997	Hs.23804	ESTs	0.82	0.15
	106605	AA457718	Hs.211103	Homo sapiens mRNA; cDNA DKFZp554B076 (fr	0.99	0.07
	106657	AA461086	Hs.16578	ESTs	1.17	0.4
	106773	AA478109	Hs.188833	ESTs	1.46	0.43
45	106797	AA478962	Hs.169943	ESTs	1.18	0.32
	106844	AA485055	Hs.158213	sperm associated antigen 6	0.98	0.51
	106870	AA487576	Hs.26530	serum deprivation response (phosphatidyl	1.05	0.14
	106954	AA496980	Hs.204038	ESTs	1.25	0.33
	107054	AA600150	Hs.14366	ESTs	1.11	0.4
50	107292	T30407	Hs.4789	ESTs; Weakly similar to oxidative-stress	1.07	2.58
	107994	AA036811	Hs.166030	ESTs	0.7	0.21
	107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	1.02	0.48
	108041	AA041552	Hs.61957	ESTs	1.44	0.51
	108087	AA045709	Hs.40545	ESTs	1.98	1
55	108382	AA074886	Hs.67726	macrophage receptor with collagenous str	1.52	0.72
	108435	AA078787	Hs.194101	ESTs	2.53	1.53
	108480	AA081093	Hs.68055	ESTs	1.56	0.48
	109252	AA194830	Hs.85944	ESTs	2.69	3.18
	109550	F01534	Hs.26981	ESTs	1.19	0.65
60	109613	F03031	Hs.27519	ESTs	1.01	0.29
	109837	H00656	Hs.29792	ESTs	0.81	0.15
	109893	H04768	Hs.30484	ESTs	1.44	0.32
	109984	H09594	Hs.10299	ESTs	0.62	0.14
	110099	H16568	Hs.23748	ESTs	1.01	0.28
65	110837	N30796	Hs.17424	ESTs; Weakly similar to semaphorin F [H.	1.1	0.22
	111247	N59825	Hs.16762	Homo sapiens mRNA; cDNA DKFZp554B2062 (f	1.26	0.26
	111341	N80935	Hs.22483	ESTs	1.57	0.52
	111510	R07856	Hs.16355	ESTs	3.95	1
	111737	R25410	Hs.9218	ESTs	0.97	0.24
70	113195	T67112		***yc20g11.s1 Stratagene lung (#937210)	1.22	0.35
	113238	T62979	Hs.189813	ESTs	2.27	0.45
	113540	T90496	Hs.16757	ESTs	1.06	0.22
	113552	T90889	Hs.16026	ESTs	1.16	0.42
	113606	T93093	Hs.17125	ESTs	1.48	0.7
75	113695	T96965	Hs.17948	ESTs	1.54	0.28
	113946	WB4753	Hs.37896	ESTs	1.79	0.72
	114251	Z39898	Hs.21948	ESTs	1.95	0.25
	114359	Z41589	Hs.153483	ESTs; Moderately similar to H1 chloride	1.42	0.13
	115230	AA278300	Hs.182980	ESTs	2.62	0.42
80	115279	AA279760	Hs.63671	ESTs	1.79	0.91
	115566	AA395083	Hs.43977	ESTs	0.86	0.2
	115965	AA446661	Hs.173233	ESTs	0.79	0.04
	116166	AA461556	Hs.202949	KIAA1102 protein	2.29	0.68
	116279	AA485073	Hs.57362	ESTs	2.27	0.78
	117023	H88157	Hs.41105	ESTs	1.36	0.16

	117209	H99959	Hs.42768	ESTs	1.46	0.48
	118901	N90719	Hs.94445	ESTs	1.51	1
	118981	N93839	Hs.39288	ESTs	1.34	0.48
5	119073	R32894	Hs.45514	v-els avian erythroblastosis virus E26 o	1.14	0.27
	119221	R98105		"y30g11.s1 Soares fetal liver spleen	1.32	0.53
	119824	W74536	Hs.184	advanced glycosylation end product-speci	1	0.19
	119861	W80715		ESTs; Moderately similar to III ALU SUB	1.83	0.45
	120041	W92775	Hs.59368	ESTs	1.23	0.55
10	120132	Z38839	Hs.125019	ESTs; Highly similar to KIAA0886 protein	0.91	0.37
	120467	AA251579	Hs.187628	ESTs	1.87	1.91
	121314	AA402799	Hs.182538	ESTs	1.3	0.31
	121643	AA417078	Hs.193767	ESTs	2.31	0.68
	121690	AA418074	Hs.110286	ESTs	1.47	0.51
15	122633	AA454080	Hs.34853	inhibitor of DNA binding 4; dominant neg	1.31	0.63
	123978	C20653	Hs.170278	ESTs	1.52	0.32
	124214	H58508	Hs.151323	ESTs	0.93	0.35
	124357	N22401		"yw37g07.s1 Morton Fetal Cochlea Homo	1.29	1
	124438	N40188	Hs.102550	ESTs	1.36	0.7
20	125167	W45560	Hs.102541	ESTs	1.46	0.69
	125174	W51835	Hs.231082	EST	3.07	3.76
	125422	AA903229	Hs.153717	ESTs	1.34	0.3
	125561	AA117667	Hs.22978	ESTs	1.89	0.63
	125831	D60888		"HUM145B09B Clontech human fetal brain	0.94	0.36
25	127002	R35380	Hs.24979	ESTs	3.02	4.06
	127307	AA369367	Hs.126712	ESTs; Weakly similar to pL2 hypotheticala	1.01	0.69
	127609	AA622559	Hs.150318	ESTs	1.21	0.32
	127959	AI302471	Hs.124292	ESTs	2.5	1
	128458	D52193	Hs.56340	ESTs	1.13	0.33
30	128624	AA479209	Hs.102647	ESTs	1.45	0.58
	128789	AA486567	Hs.105695	ESTs	1.1	0.34
	128798	AF014958	Hs.105938	chemokine (C-C motif) receptor-like 2	1.16	0.55
	128952	R51076	Hs.107361	ESTs; Highly similar to Rap2 interacting	2.04	2.4
	129057	X62466	Hs.214742	CDW52 antigen (CAMPATH-1 antigen)	1.77	0.73
35	129210	AA401654	Hs.202949	KIAA1102 protein	1.11	0.36
	129240	W24360	Hs.237868	interleukin 7 receptor	0.91	0.41
	129402	T63781		"yc21g01.s1 Stratagene lung (#937210)	1.36	0.43
	129565	X77777	Hs.198725	vasoactive intestinal peptide receptor 1	0.67	0.08
	129593	AA487015	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	1.3	0.42
40	129626	AA447410	Hs.11712	ESTs; Weakly similar to III ALU SUBFAMI	1.28	0.46
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ublqu	1.58	1
	129898	N48595	Hs.13256	ESTs	1.13	0.53
	129958	L20591	Hs.1378	annexin A3	0.81	0.31
45	130273	U59914	Hs.153863	MAD (mothers against decapentaplegic; Dr	0.59	0.22
	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	1.44	0.76
	130657	T94452	Hs.201591	ESTs	0.96	0.42
	131061	N64328	Hs.22567	ESTs; Moderately similar to HYPOTHETICAL	1.51	0.45
	131066	F09006	Hs.22588	ESTs	0.97	0.37
	131263	R38334	Hs.24950	regulator of G-protein signalling 5	2.34	2.82
50	131589	U52100	Hs.29191	epithelial membrane protein 2	1.2	0.62
	131686	AA157428	Hs.30687	Grb2-associated binder 2	0.95	0.38
	131751	H18335	Hs.31562	ESTs	1.47	0.52
	132430	T23630	Hs.258675	EST	1.86	2.09
	132476	N67192	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regl	1.73	0.58
55	132836	F09557	Hs.57929	slit (Drosophila) homolog 3	0.91	0.29
	133120	X64559	Hs.65424	tetranectin (plasminogen-binding protein	0.82	0.2
	133488	D45370	Hs.74120	adipose specific 2	1.29	0.48
	133565	H57056	Hs.204831	ESTs	2.25	0.57
	133651	U97105	Hs.173381	dihydropyrimidinase-like 2	1.65	0.62
60	133835	AA059489	Hs.76640	ESTs; Highly similar to RGC-32 [R.norveg	1.16	0.34
	133978	W73859	Hs.78061	transcription factor 21	0.79	0.27
	133985	L34857	Hs.78146	platelet/endothelial cell adhesion molec	0.99	0.28
	134299	AA487558	Hs.8135	ESTs	1.02	0.46
	134300	U81984	Hs.166082	endothelial PAS domain protein 1	0.88	0.42
65	134323	AA028976	Hs.8175	Homo sapiens mRNA; cDNA DKFZp564M0763 (f	1.19	0.27
	134343	D50583	Hs.82028	transforming growth factor; beta recepto	1.21	0.67
	134417	D87969	Hs.82921	solute carrier family 35 (CMP-static aci	1.28	1
	134561	U76421	Hs.85302	adenosine deaminase; RNA-specific; B1 (h	2.12	0.55
	134624	W67147	Hs.8700	deleted in liver cancer 1	2.35	2.74
70	134696	H88354	Hs.8861	ESTs	1.35	0.33
	134749	L10955	Hs.89485	carbonic anhydrase IV	0.89	0.2
	134786	L06139	Hs.89640	TEK tyrosine kinase; endothelial (venous	0.48	0.21
	134869	T35288	Hs.90421	ESTs; Moderately similar to III ALU SUB	2.14	2.64
	135346	M21056	Hs.992	phospholipase A2; group IB (pancreas)	0.63	0.13
75	100113	D00591	Hs.84746	Chromosome condensation 1	1	2.15
	100147	D13668	Hs.136348	Homo sapiens mRNA for osteoblast specifi	0.5	2
	100280	D42085	Hs.155314	KIAA0095 gene product	1.02	1.39
	100335	D83391	Hs.6793	platelet-activating factor acetylhydrola	1	5.58
	100360	D78335	Hs.75939	Uridine monophosphate kinase	0.91	2.04
80	100372	D79997	Hs.184339	KIAA0175 gene product	0.75	2.03
	100486	HG1112-HT1112		TIGR: ras-like protein TC4	1.09	1.93
	100559	HG2197-HT2267		"collagen, type VII, alpha 1"	0.97	3.6
	100576	HG2280-HT2386		"caltonin/alpha-CGRP, alt. transcript	1	1
	100668	HG2981-HT3938		"TIGR: CD44 (epican, alt. transcript 12	0.85	1.9
	100906	HG4716-HT5158		Guanosine 5'-Monophosphate Synthase	1.18	2.29
85	100930	HG721-HT4827		"TIGR: placental protein 14, endometrial	1	1.45

	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	0.84	2.6
	101031	J05070	Hs.151738	*Matrix metalloproteinase 9 (gelatinase	0.77	1.52
	101111	L08424	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
5	101124	L10343	Hs.112341	*Protease inhibitor 3, skin-derived (SKA	0.62	2.67
	101175	L18920	Hs.36980	*Melanoma antigen, family A, 2"	1	1
	101204	L24203	Hs.82237	Ataxia-telangiectasia group D-associated	0.74	4.1
	101431	M19888	Hs.1076	Small proline-rich protein 1B (cornifin)	0.85	2.51
	101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	0.61	8.83
10	101511	M27826	Hs.267319	Endogenous retroviral protease	1.03	1.13
	101526	M29540	Hs.220529	Carcinoembryonic antigen-related cell ad	1.07	4.61
	101548	M31328	Hs.71642	*Guanine nucleotide binding protein (G p	0.97	1.13
	101625	M57293		*Human parathyroid hormone-related pepti	1	1
	101649	M60047	Hs.1690	Heparin-binding growth factor binding pr	1	2.7
	101724	M69225	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	1	8.98
15	101748	M76482	Hs.1925	Desmoglein 3 (pemphigus vulgaris antigen	1	2.78
	101759	M80244	Hs.184601	*Solute carrier family 7 (cationic amino	1.07	2.45
	101804	M86699	Hs.169840	TTK protein kinase	1	1
	101806	M86757	Hs.112408	S100 calcium-binding protein A7 (psorias	0.74	1.76
20	101809	M86849		*Homo sapiens connexin 26 (GJB2) mRNA, c	1	7
	101845	M93426	Hs.78667	*Protein tyrosine phosphatase, receptor-	1	1
	101851	M94250	Hs.82045	Midkine (neurite growth-promoting factor	1.13	2.6
	102083	U10323	Hs.75117	*Interleukin enhancer binding factor 2,	1.03	1.61
	102154	U17760	Hs.75517	*Laminin, beta 3 (lncin (125kD), kalin	0.94	3.62
25	102193	U20758	Hs.313	secreted phosphoprotein 1 (osteopontin;	0.34	4.59
	102305	U33286	Hs.80073	chromosome segregation 1 (yeast homolog)	1.45	2.97
	102348	U37519	Hs.87539	Aldehyde dehydrogenase 8	0.52	2.25
	102581	U61145	Hs.77256	Enhancer of zeste (Drosophila) homolog 2	0.91	2.46
	102610	U65011	Hs.30743	Preferentially expressed antigen in mela	1	3.88
30	102623	U66083	Hs.37110	*Melanoma antigen, family A, 9 (MAGE-9)"	1	1
	102669	U71207	Hs.29279	Eyes absent (Drosophila) homolog 2	1	1
	102696	U74612	Hs.239	Forkhead box M1	1.06	2.77
	102829	U91618	Hs.80962	Neurotensin	1	1
	102888	X04741	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.13	2.59
35	102913	X07696	Hs.80342	keratin 15	0.7	4.72
	102915	X07820	Hs.2258	Matrix Metalloproteinase 10 (Stromolysin	1.15	3.35
	102963	X15943	Hs.37058	*Calcitonin/calcitonin-related polypepti	1	1
	103021	X53587	Hs.85266	*Integrin, beta 4"	1.38	2.34
	103036	X54925	Hs.83169	Matrix metalloprotease 1 (interstitial c	1	14.93
40	103058	X57348	Hs.184510	Stratifin	1.25	4.17
	103060	X57766	Hs.155324	matrix metalloproteinase 11 (stromelysin	1	1.72
	103119	X63629	Hs.2877	*Cadherin 3, P-cadherin (placental)"	1.16	7.38
	103206	X72755	Hs.77367	monokine induced by gamma interferon	0.71	1.48
	103242	X76342	Hs.389	*Alcohol dehydrogenase 7 (class IV), mu	1	1
45	103312	X82693	Hs.3185	*Lymphocyte antigen 6 complex, locus D;	0.92	1.28
	103478	Y07755	Hs.38991	S100 calcium-binding protein A2	1.05	5.81
	103558	Z19574	Hs.2785	keratin 17	0.65	6.68
	103576	Z26317	Hs.2631	Desmoglein 2	0.79	1.73
	103587	Z29083	Hs.82128	ST4 Oncofetal antigen	1	3.93
50	103594	Z31560	Hs.816	*SRY (sex determining region Y)-box 2, p	0.71	7.23
	103768	AA089997		*ESTs, Highly similar to Integral membra	0.99	1.8
	104158	AA454908	Hs.8127	KIAA0144 gene product	0.96	1.29
	104558	R56678	Hs.88959	Human DNA sequence from clone 967N21 on	1.23	7.23
	104689	AA010665		ESTs	0.96	2.11
55	104733	AA019498	Hs.23071	ESTs	1.18	1.88
	104906	AA055809	Hs.26802	Protein kinase domains containing protei	1.11	3.15
	104978	AA088458	Hs.19322	ESTs; Weakly similar to III ALU SUBFAM	1.64	2.89
	105012	AA116036	Hs.8329	*Homo sapiens mRNA for IIS353, complete	1.19	3.91
	105175	AA186804	Hs.25740	ESTs; Weakly similar to unknown [S.cerev	0.9	4.63
60	105263	AA227926	Hs.6682	ESTs	0.95	2.87
	105298	AA233459	Hs.26369	ESTs	1	1.13
	105312	AA233854	Hs.23348	S-phase kinase-associated protein 2 (p45	1.32	3.01
	105719	AA291644	Hs.36793	Hypothetical protein FLJ23188	1.28	2.31
	105743	AA293300	Hs.9598	ESTs	1	1
65	106012	AA411621	Hs.8895	ESTs; same as BFH67	0.94	2.04
	106231	AA429571	Hs.38002	KIAA1355 protein	1.04	1.5
	106540	AA454607	Hs.38114	Hypothetical protein FLJ11100	1.26	2.26
	106575	AA456039	Hs.105421	ESTs	1	2
	106632	AA459897	Hs.11950	GPI-anchored metastasis-associated prote	0.87	1.32
70	106727	AA465342	Hs.34045	Hypothetical protein FLJ20764	0.87	1.59
	106906	AA490237	Hs.222024	Transcription factor BMAL2 (cycle-like f	0.61	1.6
	107059	AA608545	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re	0.48	2.67
	107104	AA609785	Hs.15243	Nucleolar protein 1 (120kD)	1.01	1.44
	107151	AA621169	Hs.8687	ESTs; procollagen I-N proteinase	0.97	2.89
75	107284	S74039	Hs.291904	Accessory proteins BAP31/BAP29	1.15	3.65
	107901	AA026418	Hs.91539	ESTs	0.72	3.44
	107922	AA028028	Hs.61460	Ig superfamily receptor LNIR precursor	1	2.48
	107932	AA029317	Hs.18878	Hypothetical protein FLJ21620	1	1
	108695	AA121315	Hs.70823	KIAA1077 protein	0.91	3.53
80	108857	AA133250	Hs.62180	ESTs	1	1
	108860	AA133334	Hs.129911	ESTs	0.73	7.3
	108990	AA152296	Hs.72045	ESTs	1	1
	109166	AA179845	Hs.73625	*RAB6 interacting, kinesin-like (rabkine	1	4.55
	109424	AA227919	Hs.85962	Hyaluronan synthase 3	1	1.28
85	109665	F05012	Hs.27027	Hypothetical protein DKFZp762H1311	1.42	2
	109970	H09281	Hs.13234	ESTs	1.13	2.16

	110015	H10998	Hs.7164	A disintegrin and metalloproteinase doma	0.84	1.95
	110156	H18957	Hs.4213	ESTs	0.94	1.41
	110561	H59617	Hs.5199	HSPC150 protein similar to ubiquitin-con	0.91	3.18
5	111223	N68921	Hs.34806	ESTs; Weakly similar to neogenin [H.sapi	0.91	3.13
	111345	N89820	Hs.14559	Hypothetical protein FLJ10540	1	1.25
	111876	R38239	Hs.293246	*ESTs, Weakly similar to putative p150 [	0.83	1.27
	111902	R39191	Hs.109445	KIAA1020 protein	0.91	0.91
	112244	R51309	Hs.70823	KIAA1077 protein	0.77	3.01
10	112973	T17271		*cDNA FLJ13308 fis, clone OVARC1001436,	1	1
	112989	T23482	Hs.89981	"Diacylglycerol kinase, zeta (104kD)"	0.55	1.03
	113047	T25867	Hs.7549	ESTs	0.87	2
	113095	T40920	Hs.126733	ESTs	1	1
	113531	T90345	Hs.16740	Hypothetical protein FLJ11036	0.42	1.44
	113970	W86748	Hs.8109	ESTs	1.17	1.73
15	114346	Z41450	Hs.130489	*ATPase, aminophospholipid transporter-I	0.86	0.82
	114407	AA010188	Hs.103305	ESTs	0.8	1.88
	114471	AA028074	Hs.104613	RP42 homolog	1.06	1.34
	114509	AA043551	Hs.101799	KIAA1350 protein	1.82	2.32
	115060	AA253214	Hs.198249	*Gap junction protein, beta 5 (connexin	0.79	1.49
20	115091	AA255900	Hs.184523	KIAA0965 protein	0.72	1.92
	115123	AA256642	Hs.236894	*ESTs, High sim to LRP1_hu low density I	0.59	1.97
	115291	AA279943	Hs.122579	ESTs	1	1.25
	115506	AA292537	Hs.45207	Hypothetical protein KIAA1335	1.15	1.48
	115522	AA331393	Hs.47378	ESTs	0.5	3.29
25	115536	AA347193	Hs.62180	ESTs	1	1
	115697	AA411502	Hs.63325	Homo sapiens type II membrane serine pro	1	6.53
	115909	AA436566	Hs.59761	ESTs	1	6.98
	115978	AA447522	Hs.69517	Differentially expressed in Fanconi anem	1	2.31
	116028	AA452112	Hs.42544	thioredoxin-like	0.99	1.68
30	116107	AA456968	Hs.92030	ESTs	1.14	1.8
	116134	AA460246	Hs.50441	CGI-04 protein	1.11	1.86
	116157	AA461063	Hs.44298	Hypothetical protein	0.99	1.9
	116158	AA461187	Hs.61762	Hypoxia-inducible protein 2	0.44	0.86
	116335	AA495830	Hs.87013	*Homo sapiens cDNA FLJ10238 fis, clone H	0.62	3.89
35	116483	C14092	Hs.76118	Ubiquitin carboxyl-terminal esterase L1	1.04	2.36
	117320	N23239	Hs.211092	LUNX protein; PLUNC(palate lung & nasal	0.51	0.64
	117557	N33920	Hs.44532	Diubiquitin	1.11	2.63
	117693	N40939	Hs.112110	PTD007 protein	0.98	1.79
40	117881	N50073	Hs.260822	Butyrate-induced transcript 1	1	1.43
	118368	N64339	Hs.48956	ESTs	0.67	2.86
	118566	N68558	Hs.42824	Hypothetical protein FLJ10718	1.21	0.83
	118695	N71781	Hs.50081	KIAA1199 see CVA7.doc	0.88	1.63
	119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	1	1
45	119845	W79920	Hs.58581	G protein-coupled receptor 87	1	1
	120102	W85428	Hs.132927	*ESTs, Moderately similar to p53 regulat	1	1
	120104	W85477	Hs.180479	ESTs	0.69	3.07
	120486	AA253400	Hs.137569	Tumor protein 63 kDa with strong homolog	1.08	12.05
	120859	AA350158	Hs.1619	Achaete-scute complex (Drosophila) homol	1	1
50	120880	AA360240	Hs.97019	EST	1	1
	120948	AA397822	Hs.104650	Hypothetical protein FLJ10292	1.04	2.15
	120983	AA398209	Hs.97587	EST	1	1
	121362	AA405500	Hs.97932	Chondromodulin I precursor	1	1
	121369	AA405557	Hs.128791	CGI-09 protein	1	1.8
55	121791	AA423378	Hs.293317	*ESTs, Weakly similar to JM27 [H.sapiens	1	1
	123005	AA479726	Hs.105577	ESTs	1	1
	123044	AA481549	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	0.95	1.88
	123160	AA488687	Hs.284235	ESTs	1.59	4.98
	123479	AA599469	Hs.135056	clone RP5-850E9 on chromosome 20	1.19	1.64
60	123571	AA608956	Hs.112619	*ESTs, Weakly similar to PQ0109 Purkinje	1.03	1.14
	123829	AA620697	Hs.112208	XAGE-1 protein	1.39	2.2
	124006	D60302	Hs.108977	ESTs	1	4.85
	124059	F13673	Hs.99769	ESTs	1.49	8.62
	124960	T15366	Hs.194766	Seizure related gene 6 (mouse)-like	0.76	0.77
65	125218	W73561	Hs.110024	NADH:ubiquinone oxidoreductase MLRQ subu	1.33	1.77
	125453	R06041	Hs.18048	*Melanoma antigen, family A, 10"	0.8	1.42
	125759	AA425587	Hs.82226	Glycoprotein (transmembrane) nmb	1.52	2.26
	125972	AA434562	Hs.35406	*ESTs, Highly similar to unnamed protein	1.05	2.48
	125994	H55782	Hs.270799	EST	1	1.95
70	126396	N70192	Hs.278956	Hypothetical protein FLJ12929	1	1.35
	126645	A167942	Hs.61635	STEAP1 (Homo sapiens BAC clone RG041D11	1	2.23
	127221	A1354332	Hs.72365	ESTs	0.73	3.27
	127479	AA513722	Hs.179729	collagen; type X; alpha 1 (Schmid metaph	0.51	1.94
	128192	A1204246		KIAA1085 protein	1.8	3.16
75	128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	0.89	0.97
	128777	U46006	Hs.10526	Cysteine and glycine-rich protein 2	1	1
	128924	AA234962	Hs.26557	Plakophilin 3	1.3	2.97
	129041	H58873	Hs.169502	"Solute carrier family 2 (facilitated gl	0.84	2.04
	129099	H50398	Hs.108660	*ATP-binding cassette, sub-family C (CFT	0.87	1.04
80	129404	AA172056	Hs.111128	ESTs	1	1
	129466	L42583		*Genbank Homo sapiens keratin 6 isoform	0.72	12.67
	129605	S72493	Hs.115947	Keratin 16 (focal non-epidermolytic palm	0.92	1.5
	129628	U26727	Hs.1174	*Cyclin-dependent kinase inhibitor 2A (m	0.85	1.93
	130023	X13461	Hs.239600	Calmodulin-like 3	0.84	1.22
85	130080	X14850	Hs.147097	*H2A histone family, member X"	0.98	1.96
	130385	AA126474	Hs.155223	stanniocalcin 2	1	1

	130410	V01514	Hs.155421	Alpha-fetoprotein	0.63	0.63
	130441	U35835	Hs.301387	"Human DNA-PK mRNA, partial cds"	1.15	3.65
	130482	L32866	Hs.1578	Baculoviral IAP repeat-containing 5 (sur	1	1.88
5	130553	AA430032	Hs.252587	Pituitary tumor-transforming 1	0.92	1.96
	130577	M35410	Hs.162	Insulin-like growth factor binding prote	1.17	4.7
	130627	L23808	Hs.1695	Matrix metalloproteinase 12 (macrophage	0.69	4.05
	130800	AA223386	Hs.19574	ESTs; Weakly similar to katanin p80 subu	1.13	2.41
	130839	AA598689	Hs.21400	ESTs	0.8	0.89
10	131046	X02530	Hs.2248	INTERFERON-GAMMA INDUCED PROTEIN PRECURS	0.8	1.15
	131244	D38076	Hs.24763	RAN binding protein 1	1.13	1.85
	131877	J04088	Hs.156346	Topoisomerase (DNA) II alpha (170kD)	1	1
	131927	AA461549	Hs.34780	"Doublecortx; lissencephaly, X-linked (	0.81	0.62
	131965	W90146	Hs.35962	ESTs	0.74	3.27
15	131978	D60008	Hs.36232	KIAA0186 gene product	1	1
	132354	L05187	Hs.211913	Small proline-rich protein 1A	0.69	1.43
	132543	AA417152	Hs.5101	ESTs; Highly similar to protein regulati	0.79	4.27
	132632	N59764	Hs.5398	guanine-monophosphate synthetase	1	1.08
	132653	U31201	Hs.54451	"laminin gamma2 chain gene (LAMC2), exon	1	1
20	132659	Z75190	Hs.54481	"Low density lipoprotein receptor-relate	0.89	0.89
	132710	W93726	Hs.55279	"Serine (or cysteine) proteinase inhibit	0.84	4.41
	132758	W52432	Hs.56105	"ESTs, Weakly similar to WDNM RAT WDNM1	1.55	2.08
	132767	L05188	Hs.231622	Small proline-rich protein 2B	0.83	1.66
	132816	M74542	Hs.575	Aldehyde dehydrogenase 3	0.55	0.55
25	132990	AA458761	Hs.18387	transcription factor AP-2 alpha (activat	1	3.53
	133070	U69611	Hs.64311	"A disintegrin and metalloproteinase dom	1.16	2
	133282	U52960	Hs.286145	"SRB7 (suppressor of RNA polymerase B, y	1	2.7
	133317	AA215299	Hs.70830	U6 snRNA-associated Sm-like protein LSm7	0.95	1.42
	133370	AA156897	Hs.72157	Homo sapiens mRNA; cDNA DKFZp5641922	1.12	2.55
30	133391	X57579	Hs.727	H.sapiens activin beta-A subunit (exon 2	1.65	1.76
	133832	H03887	Hs.241305	estrogen-responsive B box protein (EBBP)	1.02	1.39
	134032	Z81326	Hs.78589	"Serine (or cysteine) proteinase inhibit	1	1
	134168	AA398908	Hs.181634	"Homo sapiens cDNA: FLJ23602 fis, clone	0.95	1.53
	134218	AA227480	Hs.80205	Pim-2 oncogene	1.36	2.48
35	134405	R67275	Hs.82772	"collagen, type XI, alpha 1***	0.76	2.86
	134453	X70883	Hs.83484	SRX (sex determining region Y)-box 4	1.89	3.78
	134470	X54942	Hs.83758	CDC28 protein kinase 2	1.82	4.11
	134645	U87459	Hs.167379	"Cancer/testis antigen (NY-ESO-1, CTAG1,	0.82	0.83
	134781	M17183	Hs.89626	Parathyroid hormone-like hormone	1	1
40	135002	U19147	Hs.272484	G antigen 6	1	1
	100040	M97935		AFFX control: STAT1	0.92	1.25
	101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	2.92	8.5
	101664	M60752	Hs.121017	H2A histone family; member A	1	1
	102025	U03911	Hs.78934	mutS (E. coli) homolog 2 (colon cancer;	0.8	1.61
45	102031	U04898	Hs.2156	RAR-related orphan receptor A	1	1
	102221	U24576		LJM domain only 4	1	1
	102270	U30255	Hs.75888	phosphogluconate dehydrogenase	1.08	1.43
	102339	U37022	Hs.95577	cyclin-dependent kinase 4	0.88	1.32
	102391	U41668	Hs.77494	deoxyguanosine kinase	1.07	1.58
50	103000	X51956	Hs.146580	enolase 2; (gamma; neuronal)	0.91	1.49
	103395	X94754	Hs.119503	methionine-tRNA synthetase	0.89	1.32
	105638	AA281599	Hs.20418	Homo sapiens mRNA for for histone H2B; c	0.91	1.25
	105726	AA292328	Hs.9754	activating transcription factor 5	0.94	1.48
	114841	AA234722	Hs.55408	ESTs; Moderately similar to CALCIUM-DEPE	0.78	1.55
55	115206	AA262491	Hs.186572	ESTs	1	1
	115906	AA436616	Hs.82302	ESTs	0.74	2.52
	119132	R49046	Hs.107911	ATP-binding cassette; sub-family B (MDR/	1.1	1.51
	124163	H30539	Hs.189838	ESTs	1	1
	126487	AA482505	Hs.184601	solute carrier family 7 (cationic amino	1.01	1.46
60	127141	AA307960	Hs.75478	KIAA0956 protein	0.85	1.4
	128034	AA905754	Hs.75103	tyrosine 3-monooxygenase/tryptophan 5-mo	1	1.18
	128609	AA234365	Hs.102456	survival of motor neuron protein interac	1	1.5
	128895	R37753	Hs.106985	ESTs	1.7	2
	130199	Z48579	Hs.172028	a disintegrin and metalloprotease domain	1	1
65	130524	U89995	Hs.159234	forkhead box E1	1	1
	133000	U24152	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	1	1
	133658	M25766	Hs.75426	secretogranin II (chromogranin C)	1	1
	135047	AA460466	Hs.93597	ESTs	1	1
	100053	M27830		AFFX control: 28S ribosomal RNA	0.88	1.53
70	100114	D00596	Hs.82962	thymidylate synthetase	0.68	1.86
	100128	D11094	Hs.81153	proteasome (prosome; macropain) 26S subu	1.29	2.03
	100154	D14657	Hs.81892	KIAA0101 gene product	0.71	4.26
	100161	D14694	Hs.77329	phosphatidylserine synthase 1	1.02	1.56
	100168	D14874	Hs.394	adrenomedullin	0.46	1.17
	100187	D17793	Hs.78183	aldo-keto reductase family 1; member C3	1	1
75	100188	D21063	Hs.57101	minichromosome maintenance deficient (S,	0.97	1.4
	100217	D26600	Hs.89545	proteasome (prosome; macropain) subunit;	1.13	1.9
	100220	D28364		"Human mRNA for annexin II, 5'UTR (seq	1.11	1.53
	100287	D43950	Hs.1600	chaperonin containing TCP1; subunit 5 (e	1.13	2.09
	100297	D49489	Hs.182429	protein disulfide isomerase-related prot	0.92	1.78
80	100330	D56716	Hs.77152	minichromosome maintenance deficient (S,	1.07	1.61
	100355	D78129		"Homo sapiens mRNA for squalene epoxid	0.96	1.87
	100364	D78586	Hs.154868	carbamoyl-phosphate synthetase 2; aspart	1.49	2.46
	100368	D79987	Hs.153479	extra spindle poles; S. cerevisiae; homo	0.59	1.32
	100398	D84557	Hs.155462	minichromosome maintenance deficient (mj	1.08	1.9
85	100438	D87448	Hs.91417	topoisomerase (DNA) II binding protein	1	2.15

	100455	D87953	Hs.75789	N-myc downstream regulated	0.91	1.48
	100491	HG1153-HT1153		Nucleoside Diphosphate Kinase Nm23-H2s	0.99	1.41
	100518	HG174-HT174		Desmoplakin I	1.28	3.17
	100528	HG1828-HT1857		***Nexin, Glia-Derived***	0.68	1.9
5	100661	HG2874-HT3018		Ribosomal Protein L39 Homolog	1.1	5.44
	100667	HG2981-HT3127		***Epican, Alt. Splice 11***	0.8	1.97
	100830	HG4074-HT4344		Rad2	1.01	2.12
	101061	K03515	Hs.944	glucose phosphate isomerase	0.91	1.79
	101131	L10838	Hs.167460	splicing factor; arginine/serine-rich 3	1.23	1.87
10	101162	L14595	Hs.174203	solute carrier family 1 (glutamate/neutral)	1.35	2.73
	101181	L19688	Hs.73798	macrophage migration inhibitory factor (	1.03	1.78
	101183	L19779	Hs.795	H2A histone family; member O	0.57	1.3
	101216	L25876	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	0.7	2.2
	101228	L27706	Hs.82916	chaperonin containing TCP1; subunit 6A (	0.99	1.99
15	101233	L29008	Hs.878	sorbitol dehydrogenase	0.82	2.11
	101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	1.2	1.91
	101332	L47276		***Homo sapiens (cell line HL-6) alpha t	0.69	2.78
	101342	L76191	Hs.182018	interleukin-1 receptor-associated kinase	1.04	1.84
	101396	M15796	Hs.78996	proliferating cell nuclear antigen	0.95	3.55
20	101423	M18391	Hs.89839	EphA1	1	1.5
	101445	M21259	Hs.1086	small nuclear ribonucleoprotein polypept	1.21	1.96
	101505	M27396	Hs.75692	asparagine synthetase	0.93	1.6
	101525	M29536	Hs.12163	eukaryotic translation initiation factor	1.19	1.93
	101535	M30448	Hs.251669	casein kinase 2; beta polypeptide	0.96	1.42
25	101607	M38690	Hs.1244	CD3 antigen (p24)	1.11	1.25
	101624	M55998		***Human alpha-1 collagen type I gene, 3	1.17	1.98
	101758	M77836	Hs.79217	pyrroline-5-carboxylate reductase 1	1.77	3.45
	101839	M93036	Hs.692	membrane component; chromosomal 4; surfa	0.71	1.45
	101853	M94362	Hs.76084	lamin B2	0.84	1.19
30	101977	S83364		***putative Rab5-interacting protein (cl	0.89	1.9
	101992	U01038	Hs.77597	polo (Drosophila)-like kinase	0.66	1.46
	102009	U02680	Hs.82643	protein tyrosine kinase 9	1.23	3.35
	102012	U03057	Hs.118400	singed (Drosophila)-like (sea urchin fas	0.85	1.88
35	102039	U05861	Hs.201957	aldo-keto reductase family 1; member C1	0.93	2.32
	102123	U14518	Hs.1594	centromere protein A (17kD)	1	4.28
	102130	U15009	Hs.1575	small nuclear ribonucleoprotein D3 polyp	0.89	1.42
	102148	U16954	Hs.75823	ALL1-fused gene from chromosome 1q	0.8	2.95
	102210	U23028	Hs.2437	eukaryotic translation initiation factor	1.01	1.34
	102220	U24389	Hs.65436	lysyl oxidase-like 1	1.15	2.34
40	102260	U28386	Hs.159557	karyopherin alpha 2 (RAG cohort 1; impor	1.14	2.69
	102330	U35451	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta	1.05	1.7
	102423	U44754	Hs.179312	small nuclear RNA activating complex; po	1.14	2.99
	102455	U48705	Hs.75582	discoidin domain receptor family; member	1.05	2.01
	102499	U51478	Hs.76941	ATPase; Na+/K+ transporting; beta 3 poly	1.27	1.92
45	102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	0.84	1.31
	102590	U62136		***Homo sapiens enterocyte differentiati	1.11	1.6
	102676	U72514	Hs.12045	putative protein	1.04	2.17
	102687	U73379	Hs.93002	ubiquitin carrier protein E2-C	0.86	2.28
	102704	U76638	Hs.54089	BRCA1 associated RING domain 1	1.12	1.63
50	102781	U83843		***Human HIV-1 Nef interacting protein (	0.9	1.39
	102784	U85658	Hs.61796	transcription factor AP-2 gamma (activat	0.98	2.16
	102827	U91327	Hs.6456	chaperonin containing TCP1; subunit 2 (b	0.96	1.62
	102935	X13482	Hs.80506	small nuclear ribonucleoprotein polypept	1.21	4.2
55	102972	X16662	Hs.87268	annexin A8	1.25	2.32
	102983	X17620	Hs.118838	non-metastatic cells 1; protein (NM23A)	1.03	1.83
	103023	X53793	Hs.117950	multifunctional polypeptide similar to S	1.58	5.44
	103038	X54941	Hs.77550	CDC28 protein kinase 1	1.32	3.79
	103075	X59543	Hs.2934	ribonucleotide reductase M1 polypeptide	1.11	2.58
60	103168	X68314	Hs.2704	glutathione peroxidase 2 (gastrointestin	0.75	3.05
	103185	X89910	Hs.74358	transmembrane protein (63kD); endoplasm	1.01	1.97
	103212	X73874	Hs.2393	phosphorylase kinase; alpha 1 (muscle)	0.95	1.72
	103223	X74801	Hs.1708	chaperonin containing TCP1; subunit 3 (g	0.97	1.77
	103260	X78416	Hs.3155	casein; alpha	1	1
65	103262	X78565	Hs.204133	hexabrachion (tanasin C; cytotoxicin)	1.23	3.09
	103330	X85373	Hs.77496	small nuclear ribonucleoprotein polypept	1.12	2.25
	103364	X90872	Hs.75854	SULT1C sulfotransferase	2.85	4.62
	103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	1	2.48
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	1	1.53
70	103404	X95586	Hs.78596	proteasome (prosome; macropain) subunit	0.92	1.53
	103437	X98260	Hs.82254	M-phase phosphoprotein 11	0.92	1.54
	103448	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	0.55	0.96
	103605	Z35402	Hs.194657	cadherin 1; E-cadherin (epithelial)	1.32	2.51
	103646	Z68228	Hs.2340	junction plakoglobin	0.88	1.28
	103658	Z74615	Hs.172928	collagen; type I; alpha 1	1.06	2.98
75	103774	AA092898	Hs.92918	ESTs; Weakly similar to R07G3.8 [C.elega	1.88	4.66
	104261	AF008442	Hs.5409	RNA polymerase I subunit	0.87	2.17
	104276	C02193	Hs.85222	ESTs; Weakly similar to R27090_2 [H.sapi	1.4	2.49
	104289	C16281	Hs.75478	KIAA0956 protein	1.15	1.68
	104434	L02870	Hs.1640	collagen; type VII; alpha 1 (epidermolys	1.04	1.49
80	104453	M19169	Hs.123114	cystatin SN	0.38	0.76
	104611	R98280	Hs.125845	ribulose-5-phosphate-3-epimerase	1.08	2.25
	104758	AA024661	Hs.7010	ESTs; Weakly similar to ACYL-COA DEHYDRO	1.14	1.65
	105114	AA156532	Hs.11801	adenosine A2b receptor pseudogene	0.91	1.38
	105132	AA159501	Hs.247280	HBV associated factor	1.08	1.7
85	105174	AA186613	Hs.34744	ESTs	0.95	2.05

	105280	AA232215	Hs.14600	ESTs	1	1.4
	105344	AA235303	Hs.8645	ESTs	0.72	2.02
	105516	AA257971	Hs.21214	ESTs	1.35	3.56
5	105621	AA280865	Hs.6375	Homo sapiens mRNA; cDNA DKFZp564K0222 (f	1.23	1.82
	105698	AA287393	Hs.15202	ESTs; Weakly similar to oligodendrocyte-	0.98	1.28
	105705	AA290767	Hs.101282	Homo sapiens mRNA; cDNA DKFZp434B102 (fr	0.92	1.32
	105724	AA292038	Hs.22934	ESTs; Weakly similar to ZINC FINGER PROT	0.99	1.41
	105782	AA350215	Hs.21580	ESTs	1	1
	105799	AA372018	Hs.24743	ESTs	1.08	1.78
10	105807	AA393803	Hs.16869	ESTs; Moderately similar to COLLAGEN ALP	0.95	1.34
	105891	AA400768	Hs.26662	ESTs; Weakly similar to tumor necrosis f	0.87	2.25
	105936	AA404338		ESTs	1.14	1.46
	106069	AA417741	Hs.29899	ESTs; Weakly similar to ZINC FINGER PROT	1	1
	106103	AA421104	Hs.12094	ESTs	1.04	1.44
15	106140	AA424524	Hs.14912	KIAA0286 protein	1.23	2.11
	106149	AA424681	Hs.256301	ESTs	0.83	1.48
	106154	AA425304	Hs.6994	ESTs	0.77	2.05
	106182	AA426609	Hs.10862	ESTs	0.74	2.23
20	106220	AA426582	Hs.32196	ESTs; Moderately similar to metargidin p	0.97	1.99
	106228	AA429280	Hs.17719	ESTs	0.99	1.54
	106318	AA436570	Hs.9605	pre-mRNA cleavage factor Im (25kD)	0.95	2.09
	106341	AA441798	Hs.5243	ESTs; Moderately similar to pIL2 hypothe	0.98	2.66
	106432	AA448850	Hs.17138	ESTs	0.95	1.93
25	106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	1	1
	106483	AA451676	Hs.30299	IGF-II mRNA-binding protein 2	1.4	2.29
	106599	AA457235	Hs.12842	ESTs; Moderately similar to non-function	1	1.82
	106611	AA458904	Hs.26267	ESTs; Weakly similar to torsinA [H.sapi	1.49	2.78
	106654	AA460449	Hs.3784	ESTs; Highly similar to phosphoserine am	1	1.4
30	107076	AA609145	Hs.21143	ESTs; Weakly similar to fos39554_1 [H.s	1.11	1.49
	107115	AA610108	Hs.27693	ESTs; Highly similar to CGI-124 protein	1	1.03
	107129	AA620553	Hs.4756	flap structure-specific endonuclease 1	1.13	3.63
	107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	1.05	2.09
	107444	W28391	Hs.5181	proliferation-associated 2G4; 38kD	1.18	1.9
35	107481	W58247	Hs.27437	Homo sapiens kinesin superfamily motor K	0.99	2.74
	107516	X56597	Hs.99853	fibrillarin	0.94	1.77
	107529	Y12065	Hs.5092	nucleolar protein (KKE/D repeat)	1.05	2.29
	107531	Y13936	Hs.17883	protein phosphatase 1G (formerly 2C); ma	1.06	1.62
	107801	AA019433	Hs.173100	ESTs	1.03	1.4
40	107957	AA031948	Hs.57548	ESTs	0.95	1.46
	108565	AA085342	Hs.1526	ATPase; Ca++ transporting; cardiac muscl	0.59	1.35
	108780	AA128561	Hs.117938	collagen; type XVII; alpha 1	1	7.63
	108828	AA131584	Hs.71435	DKFZP564O0463 protein	1.33	2.56
	109060	AA150879	Hs.241551	chloride channel; calcium activated; fam	0.87	1.42
45	109112	AA169379	Hs.72665	ESTs	1.03	2.31
	109344	AA213696	Hs.86559	poly(A)-binding protein-like 1	0.97	1.55
	109412	AA227145	Hs.209473	ESTs; Weakly similar to REGULATOR OF MIT	0.76	1.87
	110780	N23174	Hs.22891	solute carrier family 7 (cationic amino	0.9	0.95
	110958	N50550	Hs.24587	signal transduction protein (SH3 contain	1.17	2.26
50	111018	N54067	Hs.3628	mitogen-activated protein kinase kinase	1.21	1.85
	111337	N79612	Hs.16607	ESTs; Highly similar to Myosin heavy cha	1	1.45
	112305	R54822	Hs.26244	ESTs	1	1
	112401	R61279	Hs.237536	ESTs; Weakly similar to F25B5.3 [C.elega	1.24	1.64
	112853	T02843	Hs.4351	EST	1.56	1.96
55	112869	T03313	Hs.4747	dyskeratosis congenita 1; dyskerin	1.03	1.57
	112992	T23513	Hs.7147	ESTs	1	1
	113048	T25895	Hs.184008	ESTs; Weakly similar to RNA-binding prot	1.37	2.26
	113063	T32438	Hs.5027	ESTs	1	1
	113179	T55182	Hs.152571	ESTs; Highly similar to IGF-II mRNA-bind	1.33	2.7
60	113573	T91166	Hs.15990	ESTs	0.76	1.47
	113811	W44928	Hs.4878	ESTs	0.79	1.51
	114086	Z38266	Hs.12770	Homo sapiens PAC clone DJ0777023 from 7p	0.9	1.34
	114587	AA070827	Hs.180320	ESTs; Weakly similar to GOLGI 4-TRANSMEM	1.02	1.76
	114846	AA234929	Hs.44343	ESTs	1.32	2.36
65	114964	AA243873	Hs.82184	ring finger protein 3	1.1	1.84
	115047	AA252627	Hs.22554	homeo box B5	1.01	2.36
	115166	AA258409	Hs.198907	myelin protein zero-like 1	1.05	2.31
	115167	AA258421	Hs.43728	hypothetical protein	1.52	2.52
	115239	AA278650	Hs.73291	ESTs; Weakly similar to similar to the b	0.7	2.57
70	115278	AA279757	Hs.67466	ESTs; Weakly similar to BACN32G11.d [D.m	1.14	2.12
	115652	AA405098	Hs.38178	ESTs	0.82	4.67
	115875	AA433943	Hs.43946	ESTs; Weakly similar to Weak similarity	1.2	1.98
	116004	AA449122	Hs.76086	ESTs; Highly similar to small zinc finger	0.96	1.31
	116121	AA459254	Hs.48855	ESTs	0.97	1.55
75	116129	AA459956	Hs.49163	ESTs; Highly similar to putative ribonuc	1.08	2.73
	116190	AA464963	Hs.67776	ESTs	0.8	1.57
	116312	AA490494	Hs.65403	ESTs	1.37	2.65
	116732	F13779	Hs.165909	ESTs	0.92	1.8
	117602	N35020	Hs.44685	ESTs; Weakly similar to GOLIATH PROTEIN	1.15	1.84
80	117950	N51394	Hs.75478	KIAA0956 protein	1.04	2.36
	117992	N52000	Hs.172089	Homo sapiens mRNA; cDNA DKFZp566B0222 (f	0.62	1.29
	118785	N75386	Hs.111867	GLI-Kruppel family member GLI2	1	1
	119717	W69134	Hs.57987	ESTs	1	1.4
	119814	W74069	Hs.58350	ESTs	0.78	1.77
85	120128	Z38499	Hs.91448	MKP-1 like protein tyrosine phosphatase	0.86	1.46
	120242	Z98443	Hs.86366	ESTs	0.83	2.01



	120483	AA252994	Hs.1578	apoptosis inhibitor 4 (survivin)	0.74	1.64
	121054	AA396604	Hs.97387	ESTs	1.05	1.93
	121326	AA404246	Hs.97031	ESTs; Weakly similar to Similar to phyto	0.98	1.3
5	121376	AA405699	Hs.166232	ESTs; Moderately similar to SODIUM- AND	0.91	1.83
	121457	AA411448	Hs.208985	ESTs	0.91	1.59
	121780	AA422086	Hs.124660	ESTs	0.46	0.55
	121781	AA422150	Hs.98370	cytochrome P540 family member predicted	1.07	1.54
	121844	AA425732	Hs.98485	gap junction protein; beta 2; 26kD (conn	0.94	1.4
10	122059	AA431737	Hs.98749	EST	1.93	2.33
	122338	AA443311	Hs.98998	ESTs	1	1
	122354	AA443772	Hs.186692	ESTs	0.88	1.39
	122591	AA453265	Hs.99311	ESTs; Weakly similar to MRJ (Hsapiens)	2.28	2.93
	122790	AA460156	Hs.99556	ESTs	0.88	1.3
15	123398	AA521265	Hs.105514	ESTs	1	1.93
	123518	AA608531	Hs.170313	ESTs	1	1
	123673	AA609471	Hs.112712	ESTs	1	1.15
	124000	D57317	Hs.74861	activated RNA polymerase II transcriptio	0.74	1.12
	124367	N24006	Hs.99348	distal-less homeo box 5	0.67	1.1
20	124447	NA8000	Hs.140945	Homo sapiens mRNA; cDNA DKFZp586L141 (fr	1.19	1.7
	125756	W25498	Hs.81634	ATP synthase; H+ transporting; mitochond	0.93	1.59
	125769	AI382972	Hs.82128	5T4 oncofetal trophoblast glycoprotein	1.65	6.76
	125852	H09290	Hs.76550	Homo sapiens mRNA; cDNA DKFZp564B1264 (fr	0.72	2.26
	125924	AA526849	Hs.82109	syndecan 1	1.22	2.25
25	126037	M85772	Hs.6066	KIAA1112 protein	1.36	1.63
	126214	N29455	Hs.74316	desmoplakin (DPI; DPII)	1.93	3.55
	126414	N78770	Hs.223439	ESTs	1.21	1.66
	126737	AA486132	Hs.62741	ESTs	1	1
	126743	AA179253	Hs.172182	poly(A)-binding protein; cytoplasmic 1	1.3	2.16
30	126926	AA179546	Hs.832	ESTs; Highly similar to INTEGRIN BETA-8	2.53	2.8
	127432	AA501734	Hs.170311	heterogeneous nuclear ribonucleoprotein	1.57	2.12
	128218	H02582	Hs.99189	ESTs; Moderately similar to recombinatio	1.24	2.09
	128527	M31523	Hs.101047	transcription factor 3 (E2A immunoglobul	1.08	1.78
	128568	X60873	Hs.247568	adenylate kinase 3	1.23	3.48
35	128584	M11433	Hs.101850	retinol-binding protein 1; cellular	0.87	2.42
	128628	C14037	Hs.251978	EST	1.22	1.9
	128691	W27939	Hs.103834	ESTs	1.1	1.73
	128714	V00599	Hs.179661	Homo sapiens clone 24703 beta-tubulin mR	0.92	1.17
	128733	AA328993	Hs.104558	ESTs	1.34	1.94
40	128781	X85372	Hs.105465	small nuclear ribonucleoprotein polypept	0.9	1.34
	129052	AA496297	Hs.182740	ribosomal protein S11	2.59	3.19
	129095	L12350	Hs.108623	thrombospondin 2	1.04	3.2
	129241	AA435555	Hs.109706	ESTs; Moderately similar to HN1 (M.muscu	0.95	1.61
	129665	M88458	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic relic	1.28	2.63
45	129703	AA401348	Hs.179699	ESTs	0.97	1.63
	129720	AA476582	Hs.12152	ESTs; Moderately similar to SIGNAL RECOG	1.09	1.79
	129850	N20593	Hs.56845	GDP dissociation inhibitor 2	0.74	1.68
	129896	AA043021	Hs.13225	UDP-Gal:betaGlcNAc beta 1,4- galactosylt	1.43	4.19
	130069	AA055896	Hs.146428	collagen; type V; alpha 1	1.17	1.98
50	130405	H88359	Hs.155396	nuclear factor (erythroid-derived 2)-lik	1.26	1.79
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	1	1
	130599	M91670	Hs.174070	ubiquitin carrier protein	1.07	1.66
	130867	J04093	Hs.2056	UDP glycosyltransferase 1	1	4.8
	131009	AA063596	Hs.22142	ESTs; Weakly similar to NADH-CYTOCHROME	0.93	1.05
55	131028	U20240	Hs.2227	CCAAT/enhancer binding protein (C/EBP);	1	1.23
	131083	U66661	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.1	1.8
	131091	T35341	Hs.22880	ESTs; Highly similar to dipeptidyl pepti	1.28	1.98
	131144	C14412	Hs.23528	ESTs; Highly similar to HSPC038 protein	1.43	2.06
	131148	C00038	Hs.23579	ESTs	0.88	3.38
60	131164	Y00503	Hs.182265	keratin 19	1.19	2.77
	131185	M25753	Hs.23960	cyclin B1	0.86	3.84
	131219	C00476	Hs.24395	small inducible cytokine subfamily B (Cy	0.66	2.96
	131454	AA455896	Hs.2699	glypican 1	0.99	1.54
	131687	L11068	Hs.3069	heat shock 70kD protein 9B (mortalin-2)	1	1.18
65	131689	AA599653	Hs.30696	transcription factor-like 5 (basic helix	1	1.95
	131692	D50914	Hs.30736	KIAA0124 protein	1.55	2.39
	131786	AA135554	Hs.32125	ESTs	1	1.33
	131843	AA195893	Hs.184062	ESTs; Moderately similar to putative Rab	0.83	1.63
	131860	U02082	Hs.334	Oncogene TIM	1.08	2.2
70	131884	H90124	Hs.3463	ribosomal protein S23	1.23	1.24
	131903	AA481723	Hs.3436	deleted in oral cancer (mouse; homolog)	0.91	1.18
	131945	M87339	Hs.35120	replication factor C (activator 1) 4 (37	1	2.8
	131958	AA093998	Hs.3566	ESTs; Highly similar to phosphorylation	0.87	1.36
	131964	W42508	Hs.3593	ESTs	1	1.25
75	132001	J00277	Hs.37003	v-Ha-ras Harvey rat sarcoma viral oncoge	1.12	1.43
	132040	AA146843	Hs.172894	BH3 Interacting domain death agonist	1	1.55
	132065	D82226	Hs.211594	proteasome (prosome; macropain) 26S subu	0.89	1.27
	132109	AA599801	Hs.40098	ESTs	1	1.05
	132112	AA150661	Hs.40154	[jumonji] (mouse) homolog	0.99	1.44
80	132123	AA447123	Hs.250705	ESTs	1.06	2.46
	132162	H89551	Hs.41241	ESTs	1.08	2.46
	132180	AA405569	Hs.418	fibroblast activation protein; alpha; se	1.02	4.56
	132309	AA460917	Hs.2780	jun D proto-oncogene	1.16	1.8
	132371	AA235448	Hs.46677	ESTs	0.8	1.26
85	132618	AA253330	Hs.5344	adaptor-related protein complex 1; gamma	0.5	1.49
	132736	U68019	Hs.211578	MAD (mothers against decapentaplegic; Dr	1.21	1.81

	132771	AA488432	Hs.56407	phosphoserine phosphatase	1	1.3
	132833	U78525	Hs.57783	eukaryotic translation initiation factor	0.91	1.43
	132922	T23641	Hs.6066	KIAA1112 protein	1.16	1.53
5	132959	AA028103	Hs.61472	ESTs; Wealdy similar to unknown [S.cerev	1.02	1.88
	132994	AA505133	Hs.7594	solute carrier family 2 (facilitated glu	0.72	2.97
	133005	C21400	Hs.103329	KIAA0970 protein	0.88	1.34
	133065	X62535	Hs.172690	diacylglycerol kinase; alpha (80kD)	0.93	1.23
	133083	N70633	Hs.6456	chaperonin containing TCP1; subunit 2 (b	1.14	1.76
10	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso	0.97	1.43
	133134	T89703	Hs.65648	RNA binding motif protein 8	1.1	1.8
	133195	AA350744	Hs.181409	KIAA1007 protein	2.29	2.69
	133313	AA249427	Hs.70704	ESTs	1.07	1.68
	133331	T62039	Hs.158675	ribosomal protein L14	0.85	1.18
15	133438	D13370	Hs.73722	APEX nuclease (multifunctional DNA repai	0.91	1.45
	133445	T99303	Hs.73797	guanine nucleotide binding protein (G pr	0.94	1.68
	133483	X52426	Hs.74070	keratin 13	0.85	1.14
	133492	L40397	Hs.74137	transmembrane trafficking protein	1.1	1.69
	133504	W95070	Hs.74316	desmoplakin (DP1; DP11)	0.7	6.21
20	133517	X52947	Hs.74471	gap junction protein; alpha 1; 43kD (con	0.95	1.3
	133540	D78151	Hs.74619	proteasome (prosome; macropain) 26S subu	0.91	1.25
	133584	L07758	Hs.172589	nuclear phosphoprotein similar to S. cer	0.84	1.29
	133627	U09587	Hs.75280	glycyl-tRNA synthetase	1.09	1.99
	133671	T25747	Hs.75471	zinc finger protein 146	1.02	1.5
25	133859	U86782	Hs.178761	26S proteasome-associated pad1 homolog	1.11	3.33
	133865	F09315	Hs.170290	discs; large (Drosophila) homolog 5	1.84	6.7
	133913	W84712	Hs.7753	calumenin	1.15	1.86
	133963	L34587	Hs.184693	transcription elongation factor B (SIII)	1.3	1.91
	133982	U47621	Hs.207251	nucleolar autoantigen (55kD) similar to	1.3	1.99
30	134100	L07540	Hs.171075	replication factor C (activator 1) 5 (36	0.72	1.65
	134110	U41060	Hs.79136	LIV-1 protein; estrogen regulated	1.04	1.62
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	1	1.55
	134161	U97188	Hs.79440	IGF-II mRNA-binding protein 3	0.82	1.95
	134193	F09570	Hs.7980	ESTs	0.98	1.48
35	134367	X54199	Hs.82285	phosphoribosylglycinamide formyltransfer	1	2.8
	134402	U25165	Hs.82712	fragile X mental retardation; autosomal	1.26	2
	134457	D86963	Hs.174044	dishevelled 3 (homologous to Drosophila	1	1.47
	134469	X17567	Hs.83753	small nuclear ribonucleoprotein polypept	0.94	1.57
	134498	M63180	Hs.84131	threonyl-tRNA synthetase	1.2	2.64
40	134501	W84870	Hs.211568	eukaryotic translation initiation factor	0.84	1.36
	134507	M63488	Hs.84318	replication protein A1 (70kD)	1.7	2.93
	134548	U41515	Hs.85215	Deleted in split-hand/split-toot 1 regio	1.48	2.73
	134599	X99226	Hs.86297	Fanconi anemia; complementation group A	1.36	2.22
	134692	R73587	Hs.8850	a disintegrin and metalloproteinase doma	0.77	1.64
45	134693	N70361	Hs.8854	ESTs	1.09	1.82
	134806	Z49099	Hs.89718	spermine synthase	0.98	1.35
	134821	Z34974	Hs.198382	plakophilin 1 (ectodermal dysplasia/skin	0.99	1.4
	134864	Y08999	Hs.90370	actin related protein 2/3 complex; subun	0.95	1.42
	134914	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.16	1.29
50	134953	L10678	Hs.91747	profilin 2	0.95	1.76
	134993	AA282343	Hs.9242	purine-rich element binding protein B	0.98	1.73
	135051	C15324	Hs.93668	ESTs	1.35	2.11
	135158	U51711		Human desmocollin-2 mRNA; 3' UTR	0.86	1.16

Table 1B shows the accession numbers for those keys in Table 1A lacking unigenal D's. For each probaset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the Accession column.

Pkey: Unique Eos probaset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT	Accessions
65	100661 23182_1 100667 26401_3	BE623001 L05096 AA383604 AW966416 N53295 AA460213 AW571519 AA603655 L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700 AW606203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066 A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153 BE546944 T69231 AW377441 AA907406 H50799 AW051416 A1420712 BE620922 A1279161 AA992549 W47198 BE005241 A1342696 H50700 A1959974 A1863855 AA374490 AW130675 A1950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 A1673068 A1887890 AW804171 A1675961 AW804172 AA778841 AL048050 A1127757 A1095568 AW204965 AW468978 W31898 A1052595 A1278771 BE464018 A1081503 A1824196 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 A1280942 T27619 BE621435 N66010 AW589527 A1160414 AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444 AA054555
75	100668 26401_3	L05424 X56794 S66400 X55150 W60071 AW351820 X55938 M83326 BE005289 BE070059 M83324 BE005248 BE069717 BE181648 BE069700 AW606203 BE069721 AW382138 AW803776 BE463954 BE005334 BE005274 T27386 AA932714 AA972695 AW377728 A1632506 T29066 A1783934 AW377727 BE163715 AL047291 AA279047 AA523003 BE008048 BE440141 W23614 BE090519 BE092193 N29181 N20358 N44153 BE546944 T69231 AW377441 AA907406 H50799 AW051416 A1420712 BE620922 A1279161 AA992549 W47198 BE005241 A1342696 H50700 A1959974 A1863855 AA374490 AW130675 A1950633 AA146687 H99482 X55150 BE005414 BE005339 N28294 A1673068 A1887890 AW804171 A1675961 AW804172 AA778841 AL048050 A1127757 A1095568 AW204965 AW468978 W31898 A1052595 A1278771 BE464018 A1081503 A1824196 AA513211 AA411062 AW084376 N48752 AA703209 N35580 AW059918 AA054563 A1280942 T27619 BE621435 N66010 AW589527 A1160414 AA283090 AA962536 H82726 W52115 W45432 W60433 AA577548 AA146714 BE150994 AA054615 AW796025 AW382768 BE565671 C00444 AA054555
80		J04088 NM_001067 AF071747 AJ011741 N85424 AL042407 AA218572 BE296748 BE083981 AL040877 AW499918 AW675045 A17813 BE081283 AA670403 AW504327 BE094229 AA104024 A1471482 A1970337 AA737616 A1827444 AW003286 A1742333 A1344044 A1765634
85	101332 25130_1	

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

100780 458\_127  
100830 4002\_1  
100906 4312\_1  
100930 16665\_1  
102221 3861\_1  
101809 32963\_1  
102590 15532\_1  
101977 29073\_1  
102781 20812\_1

AW172827 AA095289 BE046383 A1734240 W16899 A1660329 A1289433 AA933778 AW469242 AA468838 AA806983  
AA625873 W76031 BE206307 AA550803 A1743147 A1990075 AA948274 AA129533 A1635399 AA605313 A1624659 AW594319 A1221834 A1373434  
AA307706 BE550282 A1760467 A1630636 A1221521 AW674314 AW078889 A1937332 A1686699 A186928 AW074595 A127486 A1076544  
A1910815 H17814 AA310903 AW137854 T19279 AA026682 AA306035 AW383390 AW383389 AW383422 AW383427 AW383395 H09577  
AA306247 AA352501 AW403639 F05421 AA224473 AA305321 H93904 AA089612 AW391543 AW402915 AW173382 AW402701 AW403113  
R94438 N73126 H93466 AA090928 AA095051 T29025 AW951071 L47277 L47276 A1375913 BE384156 W24652 AA746288 AA568223 BE090591  
H93033 N57027 AA504348 AA327653 AW959913 N53767 AA843715 A1453437 AW263710 A1076594 AA583483 AW873194 AW575165 A128799  
A1803319 AL042776 AW074313 A1887722 A1032284 AA447521 A123885 N29334 A1354911 AW090687 AA236763 AA435535 AA236910  
AA047124 AA236734 AW514610 H93467 AA962007 A1446783 AA127259 A1613495 A1686720 A1587374 AA936731 AA702453 A1859757  
AA216786 A1251819 A1469227 AA806022 A1092324 N71868 AA968782 AA236919 AA809450 AA227220 AA765284 A192007 AA768810  
AA805794 AA729280 AA806238 AW768817 N71879 A1050686 AA505822 AA668974 A1688160 BE045915 AW466315 AA731314 AA649558  
AA834316 AW591901 AW063876 AW294770 A1300266 A1336094 A1560380 AA721755 H09978 D20305 D29155 AW821790 BE150664 F01675  
A1457474 AW466316 AA550969 AA630788  
BE561958 BE561728 BE397612 BE514391 BE269037 BE514207 BE562381 BE514256 BE514403 BE514250 BE397832 BE269598 BE559866  
BE396881 BE560031 BE514199 BE560037 BE560454  
AC004770 W05005 AA356068 AA094281 H29358 T56781 AW875313 L37374 BE312466 BE311755 BE207106 BE293320 BE018115 AW239090  
BE548830 AW247547 AA776062 BE397382 AA486713 T10111 T09340 AW498961 BE547280 AA356003 AW581520 AW875331 AA580720  
AW875336 BE276873 BE408229 AW188148 BE255166 BE253761 AW793727 AW373141 AW581548 AA471223 AA305950 BE263976 AA626820  
BE257409 AW309662 AA090655 C00312 BE312741 BE407213 AA209352 AW298199 AW248553 AW297794 AW731722 BE300586 AW731972  
AW615446 BE301599 AW615520 AA486714 AW440257 AA196516 AA564630 AA618079 AW192592 AW474985 AA604580 BE27461 AA766440  
A1680394 A135548 A1683224 A1581126 AW245096 AW194154 H29274 N70363 AA629758 AA580602 AA862006 A1863841 A1097667 A1928583  
A1358774 BE243487 AA620553 AA653297 AA292690 T10110 T38906 AA908544 AA340930 A185438 T03328 T28844 A1687010 A1664965  
A1872575 BE388740 T56780 AW373138 BE258717 AA699571  
ALJ076916 BE298110 AW239395 AW672700 NM\_003875 U10860 AW651755 BE297958 C03806 A195876 AA644165 T36030 AW392852  
AA446421 AW881866 A1459428 BE548103 T96204 R44457 N78225 A1564549 AW004984 AW780423 AW675448 AW087890 AA971454 AA305698  
AA879433 AA350609 A1394371 AA928053 A1378367 N59764 A1364000 A31285 T81090 AW674657 AW674987 AA879396 AW673412 BE063175  
AW674408 A1202011 R00723 A153769 A1460161 AW079585 AW275744 A1873729 D25791 BE537646 T81139 R00722  
J04129 NM\_002571 AA293088 AA477016 AA404631 T28299 AA476904 AA433955 AA430486 AA495907 A1511391 AA291495 AA402723 W25651  
AW706801 T34489 D56106 D56351 A1936579 AW023219 AW889335 AW889120 AW889232 AW889175 BE093702 AW889349 AA147546  
A1804952 AA479874 A1597661 A1420511 AA479738 AA421417 AA421247 AA436220 ALD47797 M34046 N42277 AA228076 W02698 AA420297  
AA434011 A1369971 AA479731 A1855541 A1418020 AA421246 AA452764 AL048051  
NM\_006769 U24576 AW161961 AW160473 AW160465 AW160472 AW161069 A1824831 AW162635 A1990356 AW162477 AW162571 A1520836  
AW162352 AW162351 AW162752 A1962216 A1537346 AA853902 H17667 BE045346 BE559802 BE255391 AA985217 AA235051 A129757  
AW366451 T34489 D56106 D56351 A1936579 AW023219 AW889335 AW889120 AW889232 AW889175 BE093702 AW889349 AA147546  
A1952998 AA912579 A1433556 AW902211 R64717 AW157236 A1815242 D45274 AW263991 AA442920 AA129965 A1035713 A1932325 A1949082  
A142826 A1684160 A1701987 A1678954 A1827349 BE463635 AW628092 AW302281 AA493203 BE348856 BE536419 AW193969 AW673561  
AW592609 A1224044 H43943 AA091912 R49632 R48353 A1568409 R48256 A1188046 H27986 H43899 A1678759 A1680310 A1624220 H17052  
AA156410 N56082 A1899430 AA664529 T09406 T10459 AA627506 A1379684 N83831 N88633 AW022651 AA971281 AA248036 A1039197  
A1914689 AA973825 ALD47305 AA129966 A1798369 AW264348 A1445879 A1658759 N67824 A1933507 A1216121 A1333174 T10972 A1376028  
A186756 A1273778 AA610487 A1797946 AA853903 AA903939 A1338587 A1278494 AW627595 AA904019  
NM6849 AA315280 NM\_004004 AA315269 BE142653 AA461400 AW802042 BE152893 AW383155 AA490688 AW117930 AW384563 AW384544  
AW384566 AW378307 AW378323 AW390805 AA257102 AW378317 AW276060 AW271245 AW378298 AW384497 A1598114 AW264544 A1018136  
AW021810 AA961504 AW086214 AW771489 AW192483 A1290266 AW192488 AW384490 AW007451 AW890895 AA554460 AA613715  
AW020651 A1783695 A1589498 A1917637 AW264471 AW384491 A1816732 AW368530 AW368521 AW368463 AA461087 A1341438 A1970613  
A1040737 A418400 AA947181 AA962715 A1280695 AW769275 AW023591 A1160977 AA055400 N71882 AA490466 AW243772 AW316636  
A1076554 AW511702 N69323 H88912 AA257017 A1952506 H88913 A1912481 AA600714 BE465701 N64149 C00523 N64240 AA677120  
RG1573 BE005029 X98091 AA297307 BE537267 BE566138 BE566139 F11561 BE564795 BE588776 AW064005 BE588479 BE380035 BE567012  
BE566834 BE566568 AA298090 BE566043 BE566813 BE566818 AA283070 BE565414 BE566738 BE566855 BE566567 BE566116 BE566433  
U62136 AF049140 BE567057 BE567287 BE567403 BE564316 BE567400 BE568854 BE566588 AA487772 AA071363 AW732642 BE564996  
AA297763 AA278550 AA421083 AA298184 AA091007 AA984577 AA205916 N28759 AL031291 C15757 C15761 H02728 BE566410 AA129335  
AA419499 N87741 BE379689 BE004824 BE379611 D25874 AA148454 AA323654 AW950311 AA448795 AW749423 AA773386 AA773843  
AW023027 BE348580 BE504258 BE549990 BE220200 A1073334 A1020679 AA975515 D61421 A1168688 AA102843 AW246821 A1276203  
A1074054 A1633624 A162927 A148926 N50969 A130891 AA410994 AW373025 AA148455 H02620 AA688293 A1246318 N22220 A1917777  
A1050641 A1927286 AA663794 AW368662 AW627826 AW078734 A1253080 AA749154 AA832236 A1192358 AW024676 AA448676 AA74891  
BE439487 AA661534 AA258051 A1090546 AA955157 A1051011 AA584421 A1026032 AW591338 AW589563 AA776914 AW024684 AA421002  
F09219 BE464500 A1383595 AA954244 AA601583 AA737304 AA195549 AA805778 A1055875 AA164942 AW013961 A1572608 AW514211 D59441  
AW582574 AA160935 BE566501 BE564612 BE565353 BE568195 BE565447 BE568302 BE566097 BE565470 BE584249 AL036217 AW749424  
BE567494 AA102842 A14761 AV661237 C14211 AA651866 AW798997 AA470805  
AF112213 AL050318 T24804 AW248136 BE385341 BE263177 W16677 BE250224 BE563669 BE267405 BE546577 AV651354 AV651292  
A1346903 A1539128 A1189171 S83364 AW073849 A1816760 AW073309 A1422690 AA296692 A1860301 A1805446 N77735 A1340328 BE092530  
AW028742 BE088442 AA657742 AA742438 AW170086 A1036920 A432379 N36073 A1936194 AA868655 AA983612 A1077505 BE080433  
A1375014 A1126547 A1348244 A1346077 A1748952 N26915 A1753574 A1093341 A1278762 BE092517 N74204 H06158 T58149 A129303 N58366  
AA524456 BE122661 AA542925 A1246120 A1735203 AA706829 AA877544 A1082289 AA926687 N92840 AW249758 AA934763 AW998363  
A1286332 N25202 A1240209 AW118892 N80744 R35655 A1342321 A1340141 AW878792 A1857321 H09610 W04601 AW006650 AA125006  
AA553675 A1052791 AW059835 A1041806 AA814658 AW002059 AA729483 A1609301 AA94633 AA903651 A1459183 T56072 AW088630  
AA126112 A1800091 A1561215 H17502 AW475072 A1819003 A1683272 A1262701 AW793140 T81787 R99586 A1275160 A1310420 A1698929  
AA159174 A1827968 F30305 F30309 AA806662 A1091923 AW878722 AA583430 AW571913 A1674584 AA292533 A1079471 AA642325 AA719050  
AW793172 AA305476 AW103745 T23459 N79525 A1784438 AA534551 AW193751 A1074360 BE281214 T32229 W25056 W01205 T63086  
AW795348 A1361287 AW795353 AW795349 AA594759 A1400295 D11489 A1370689 AA482356 AA485295 W40151 AA564661 AW300745  
A1346938 A1374975 A422782 AW193899 AA612604 A1183409 AA998156 AW366963 AW366977 A1284860 AA846503 A1985064 AA844576  
AA737921 AA873274 BE241546 BE241540 AA484058 AW468970 AA127876 AA159120 AW001568 AW795213 AW795258 AW795330 BE250589  
BE387572 AA910895 AA161217 BE250380 W31500 T95167 A1719306 A1359224  
BE258778 BE281230 BE410044 T33723 AW672694 AW410439 NM\_006429 AF026292 T35505 BE542333 T08940 AU076737 AW247471  
BE393215 AW328640 BE542408 T32170 BE302544 T31955 BE206898 BE275738 T32570 BE386426 BE298746 BE389937 BE293991 BE316289  
BE389578 R34739 R15312 BE279365 BE277756 AL036019 T33725 BE277779 BE302962 AL047294 BE276505 T09070 T33673 BE312580  
AW387774 BE257175 AW674367 BE253331 BE270344 BE299831 BE273576 T32062 A1751831 BE618381 AA304899 BE252268 U46364  
BE256709 BE207199 BE256209 BE251941 BE250791 BE313955 BE269806 BE543623 BE279212 BE252289 T31699 BE262220 T31669  
AA315781 AA192212 N84547 BE292737 BE259631 AA232179 A133144 T31292 AA315945 BE407301 BE251184 BE409006 A180158 A1904003  
A1904114 AW651768 AW651763 R52747 BE271897 U83843 C05298 BE261609 BE255973 AA351650 N48631 BE263637 AW452910 AA328465  
AA324549 AW579525 BE252256 BE257551 ALD48332 BE208630 AA359336 AW327897 AA151742 AA305816 BE076862 BE076796 BE263161  
AA323785 AA676588 AA626565 AA078917 W87657 R09002 R94021 AA312032 BE276665 AA295608 AW407162 AA329374 AW877912 N27885  
AA369256 AA369058 BE250476 N85427 BE265569 A1278639 A1816576 A1691037 AW328583 A1567949 A1983455 A1927732 AA811297 A1571508  
AW073674 BE259639 BE467326 A1828796 A1816578 AW511604 A1921213 AW152427 A1795787 A1801618 AW168866 A1628144 A1890339  
AW173890 AW511540 BE535620 AA383014 BE301164 A1866596 AW514909 AA658050 AW575243 AA074631 A1093488 AW575408 AW675443  
AW615636 AW732207 AW377638 AA321784 AA641629 AA633105 AA527640 AW129146 AW615572 BE294607 AA483902 AW475032 BE378532

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

AA872808 AI469368 AW105268 BED47301 AW591843 AW410066 AW517153 AI950495 AA746641 AI94878 AA873185 AI696911 AA548825  
AA911505 AA148762 AW674535 AI587329 BE328328 AW270348 AA158225 AW117705 AW474997 AW519193 AA614757 AW664383 AI082647  
AW590973 AI476711 AA192213 N88741 BE464552 AW072679 AI453708 AA152166 AA805924 AI581078 AI125768 AW173484 AI961980  
BE300766 AI195988 AI636792 AW247333 AW272851 AA078818 AA150012 AA551232 AA678821 AW873869 AW768266 AI660315 AA192120  
AA814551 AA157994 AA318886 AI582952 AW089224 AI356098 AI343694 AW072598 N21054 AI301249 AA474294 H17917 AW328534  
AW248898 AI751830 AA907816 R08898 AW087989 AI828300 AA148596 AI269577 T33426 AA213571 AI973201 AA666279 R49612 AI573183  
AW799762 AW391008 AW769666 AI962097 AA75204 D57490 AW517531 BE245270 AW470008 T33427 AW005731 AI795795 T23753 AW272981  
T15747 AA552875 T23644 AW361289 AI758558 BE207435 AA876958 T03361 AA883569 F37533 AA582321 AW082524 R42212 AA973847  
T18900 AA06202 AI559867 AI302418 AA948667 AA745670 T08939 T33724 T33722 BE621568 D57489 D25806 BE621151 F16510 C05966  
T35127 AA630427 AI933481 AA309426 AI918440 BE561854 BE618866 BE394675 BE296173 AW951687 BE383739 BE616141 BE312730  
BE535351 AW080575 BE313330 BE618664 AI354390 AA847315 BE544509 BE15212 BE297833 BE278808 BE544844 AW090178 AI890664  
BE546708 AW189943 BE274412 BE382399 BE266392 BE254949 BE280696 BE383237 BE261756 BE257721 BE312683 BE275476 BE514880  
BE545314 BE313587 BE384537 BE386691 BE264813 AW592575 AI363332 AI278641 AI795791 BE222662 AW249316 AA314361 AL036012  
AW402923 BE266845 AA075945 AA314436 BE384640 AW731769 AW957077 AA552234 AA573560 AW367038 AA313389 AI983873 BE410159  
BE263803 BE514339 BE409073 BE281296 BE543396 BE395387 BE088360 BE546946 BE546570 BE390626 AA074638 AA301821 AW045230  
AW552379 AI949222 AW029572 AA515843 AW272394 BE250234  
C14322 W74050 AI074232 AA595624 BE048955 AI148417 AI583145 AI473460 AI801688 AW573593 AI950741 AI628140 AW467921 R98105  
AI149258 AI247584 AI078378 AI139850 AA489411 W24744 R98104 AI033826 AA695989 AI033120 N55544 W88984 AW970771 AA703362  
AA099138 AA705792 AA046150 H98881 AI916674 AA953018 AI972749 AI921343 AA909044 AA094751 AI203124 AA582143 AA446654  
AW235415 R70377 AA099236 F20703 AA524436 R69484  
H04043 D60988 D60337  
AI204246 AI204250 AI194050  
H83265 T63524 AA304359 AW960551 AI672874 AI749427 AA227777 AW027055 AA971834 T49644 T54122 AI983239 AI808233 T91264 T96544  
AI350945 AI709114 R72382 T48788 R48726 AW385418 AI095484 T49645 AA928653 AA570082 AW007545 T57178 AA516413 AI93118 T57112  
AA564433 AA774503 AA357671 T59757  
W78916 AI720808 AI633854 AI632086 AI686663 N70694 AW571809 AI383592 AI201348 W80715 N91880 AW963101 AA339011  
AB033023 BE391905 BE275965 BE277872 BE003682 AA333774 BE019159 BE298024 BE299727 BE300011 BE390277 BE394764 N87550  
BE409419 BE408652 BE408197 AL119332 AA622427 AI816265 AA601118 T07318 AA019839 AA634430 BE205794 BE049461 AI042322  
AI652711 AI917645 AA630045 AW191989 AI817882 T17271 AI803663 AI095533 H46019 AW592438 AI624836 AI675552 D51149 AW132058  
AA639614 AI925762 AW088153 T17455 AA018640 AW751475 BE300241 AI816255 BE391981 AW408671 AA363910 AW875446 AW875703  
AW875926 AW875645 AW875647 AW938037 AL138042 AW892619 BE243018 AW995454 BE246381 BE090982 BE278921 AW957842 AA262454  
H30121  
W72062 AF088057 W76255 AI827219 AI631461 AW449295 AI354957 AI913803 T62772 AI222040 T62921 T63781  
AI678765 H12175 R14664 AI914049 AA995383 H08009 H19418 AW953728 AI358021 AA587361 AI269377 AA369905 AW957113 H27693  
AI300474 H73776 W74397 AA579604 AI131018 W72331 AI719085 AA568348 AI859045 AI814819 AI888714 BE467470 AW131268 H19419  
H27694 AI342165 AI914155 AA534872 BE018176 R60206 H11647 R45641 AI860466 BE301656 AI125453 AI498120 AA593735 AA879110  
AI016404 T35018 AA588397 AW449767 AA470365 BE501139 AA588354 AI337500 AW078532 Z41279 AI125449 AA935725 AA404338  
L42583 NM\_005554 L42601 BE183076 AI541221 BE140567 L42610 V01516 J00269 AW275792 AW383052 AW380143 AI541102 BE612846  
AI541344 AW128368 BE613405 BE615705 BE615530 BE615301 AW379823 AW794706 AA194806 AA194992 AW384024 AW384000 AA641239  
AI246504 AI540333 AW238581 AA640939 AI540863 AI608860 AW862564 AW366725 AW368933 AW366870 AA596020 AW794721 AW794511  
AI591181 BE182523 AW794644 AW794620 AI935234 AI068903 AI608623 AW797080 AW084935 BE182517 BE182319 AI890082 AW238346  
AW797012 BE182522 AW794838 AI608794 AW304289 AA147193 AA595995 AW381128 AW366720 AA583718 AI828416 BE122864 AW368343  
AA431080 AW082039 AW380976 AA587144 AA443636 AW872937 AW794448 AW378382 AW085761 AW794718 AW263895 AA583587  
AA583991 AA583994 AA586886 AA586880 AW368365 AI814460 AA586991 AI262829 AW378406 AA586721 AI609242 AA431973 AA232959  
AI831095 AW263854 AW378391 AW378415 AW378381 AA036990 AW238395 AI285446 BE208219 BE049526 AA583605 AA583918 AW366711  
AI285580 AW082642 AI285712 AA582875 AW591216 AW368719 AW378408 BE122835 AA582976 BE350422 AA418328 AI541454 AI565930  
AA583700 AA180575 AW238427 AI287474 AA912658 AA584223 AW238528 C17918 AW136169 AA159847 AI923797 AI609009 BE182479  
AI915198 AW378114 AA147179 AA584239 AA150532 AW168862 AW085999 AW082480 AA659742 AW079703 AI872793 AA583981 AI824571  
BE182316 BE182507 AW233331 AI824572 AI540586 D29492 BE182931 AA036948 BE551821 D29401 AW378365 C00141 D29181 D29567  
AW103359 W95238 AI991663 AA587298 BE184608 AA099833 W95121 W95150 D29584 AI934111 D29456 D29533 AW265380 D29290  
AW238463 AI212041 D29204 AA595925 D29441 AW081840 AA587018 D29323 AA582891 BE182433 BE182437 BE158295 BE182434  
AW015534 AA314369 AA250715 BE568683 AW626494 D28364 AW955678  
AI907114 AA580734 AI041945 AA101515 AA112134 D78130 NM\_003129 AA341650 T84166 AF098865 AA130976 BE089553  
T66122 AW175590 F05344 AI114790 T12900 AA194871 AA132298 D78129 AA132213 AW948930 AW948919 AA263053 AW945693 AW948840  
AA278558 R50895 N26940 N40818 AW021255 AA054851 AA563379 AW948795 AW948893 AA400356 AW948911 N85024 W78844 AI341546  
AI760182 AA286783 BE617763 BE617263 AW263690 BE049454 BE617928 AW515038 AW950584 AA601009 AI079194 AA147204 AW083163  
AA130581 AI218369 AA604784 AI808257 AI595956 AA232318 AA258065 AI471982 AA687949 AI143944 N30172 AA400196 AI769049 AI084342  
AI221380 AA484849 AI802469 H05720 AA113270 AA158138 AA076231 AI521024 AI810962 AI133616 AA805106 AA101516 R40052 R50778  
R43280 T65036 AW131924 AA114251 AA152331 F09550 AA580614 AA558927 C76491 Z38352 AA954595 C75606 W80742  
D56165 M36981 X58965 NM\_002512 BE379177 AA314836 BE256445 BE252016 AW248433 AI720933 AW085701 BE386050 BE619742  
AA482961 AI003658 H43261 AA657978 AI735072 R83138 AA722002 AA626271 AW273877 BE464626 AA071483 AA429973 AA494342  
AA620436 AA775597 AW75601 AA826847 AI192585 AA826359 AA411159 AI193419 AI204013 AA705323 AA716255 AI784611 AI081144  
AI128227 AA828464 AI148911 AI493446 AI626084 AI189180 AI721196 AI190618 AA284987 AI128543 AA632064 AI333073 AI278470 AI131688  
AI491768 AA937581 AA630065 AA834257 AW249841 AA533742 AI309756 AA961676 AI760860 AA567818 AA954238 H43655 AI302564  
AA127545 AI609219 H20426 AI042292 AI056466 AA581836 W47002 AA422057 AA937673 F29757 AA829208 AW327462 AA372098 W02144  
AA036805 AA487365 AA961037 AI139946 AA487250 AA737118 AI952504 AI242293 AA650552 AI708401 AI633133 AA630848 AA654317 F24128  
AI434165 W46252 AW043879 AI033763 F37228 AA687809 N49087 AA876981 AA506947 AI914572 AI833284 F22253 AA026222 R50166  
AI219267 N27095 AA496512 AI784222 AI289904 AA513146 AA528647 AA418700 F36721 AI880700 AI801170 AI862851 AI708633 AA524499  
AA642220 AA498628 AI718709 W80579 AI720547 F20718 AA649943 AA588229 N40503 H46029 BE262669 BE391069 BE537538 AI510751  
AI906968 AI318611 H45099 AI472604 T60667 AA373087 W32479 AA514034 BE619183 AA134672 AA127544 H26942 BE536689 AW327461  
AA422139 AW262357 AW327348 F33510 AI630382 AW827126 F27133 AI335189 AW517599 W80471 AA885814 N89681 BE393173 AA617760  
AA584268 AA460537 AA446261 H20425 N64040 AW276801 AA316367 AA071232 BE545409 AA308292 BE274447 AA380861 AA340038  
AA341806 AA865579 AI018634 AI766314 AI919302 AA872367 AA991404 AI906961 AA888375 BE621012 AA505388 AA935192 AA290828  
R50220 H08014 H44721 AW951723 AA514756 AA418708 AW673377 AA379622 AA977995 AA708224 AA708216 AI318249 AI318233 AA411160  
AA026221 AA316774 AA486908 AI500094 AA096362 AW583742 BE536422 BE618653 R70203 AA131732 AA345048 BE562720 T28342  
NM\_004415 AL031058 M77830 BE149760 AW752599 AW848723 AW376697 AW376817 AW376699 AW848371 AW376782 AW848789  
AW361413 AW849074 AW997139 AW799304 AW799309 BE077020 BE077017 BE185187 AW997196 BE156621 BE179915 BE066561 BE143155  
AW890985 BE002107 AW103521 AA857316 AW383133 BE011378 AW170253 BE185750 AW886475 BE160433 J05211 BE082576 BE082584  
BE004047 AW507238 AW377700 AW377699 BE082528 BE082505 BE082507 BE082514 AW178000 AW177933 AI905936 AW747877 AW748114  
BE148516 AW265328 AW847678 AW847688 AW365151 AW365148 AW365153 AW365156 AW365157 AW365154 AW069840  
BE000575 AW365145 BE001925 BE182166 BE144243 BE001923 AI951766 AI434518 BE184930 BE184933 AI284090 BE184941 AW804674  
BE184924 C04715 W39488 AW995615 BE184948 BE159646 AW066653 AA098991 AA131128 AA337270 AA340777 AW394371 AA582212  
R58704 AW366566 AW364859 AA025851 AA025852 AA455100 AA719958 AW352220 AW996245 BE165351 BE073457 AA377127 AW890264  
AW509750 AW391912 AW849690 T87267 AW853812 AA852213 W74149 BE009090 AA066401 H91011 AW368529 AW390272 C18467

5

10

15

20

25

AW874920 N57176 AA026480 AW576767 H93284 AA026863 AW177787 AA026654 AW177786 BE092134 BE092137 BE092136 AW177784  
AI022862 BE091653 AW376811 AW848592 AA040018 BE185331 BE182164 AA368564 AW951576 T29918 AA131077 W95048 W25458  
AW205789 H90899 N29754 W32490 R20904 BE167181 BE167165 N84767 H27408 H30146 AI190590 C03378 AI554403 AI205263 AA128470  
AI392926 AF139065 AW370813 AW370827 AW798417 AW798780 AW798883 AW798569 R33557 AA149190 C03029 AW177783 AA088866  
AW370829 AA247685 BE002273 AI760818 AI439101 AW879451 AI700963 AA451923 AI340326 AI590975 T48793 AI568096 AI142882 AA039975  
AI470146 AA946936 BE067737 BE067786 W19287 AA644381 AA702424 AI417612 AI305554 AI686869 AI568892 AW190555 AI571075 AI220573  
AA056527 AI471874 AI304772 AW517828 AI915596 AI627383 AI270345 AW021347 AW166807 AW105614 AI346078 AA552300 W95070  
AI494059 AI911702 AA149191 AA026864 AI830049 AI887258 AW780435 AI910434 AI819984 AI858282 AI078449 AI025932 AI860584 AI635878  
AA026047 AA703232 D12062 AW192085 AA658154 AW514597 AW591892 T87181 AA782066 AW243815 AW150038 AW268383 AW004633  
AI927207 AA782109 AW473233 AI804485 AW169216 AI572669 AA602182 AW015480 AW771865 AI270027 AA961816 AA283207 AI078952  
AI498487 AI348053 AI783914 H44405 AW799118 AA128330 AA515500 AA918281 W02156 AI905927 AA022701 W38382 R20795 T77861  
AW860878  
100528 45979\_1 BE386801 AU077299 AA143755 BE302747 AA853375 U30162 BE274163 BE277479 BE408180 BE274874 C15000 AA047476 N27099 AI359165  
AI638794 AI151283 AI863925 AW444977 AI207392 AA931263 AA443112 R40138 AW068638 AA351008 AA676972 R62503 AA916492 AW001865  
H42334 H38280 AA121497 AA114137 AI750938 M17783 AA383785 BE274462 AI753182 C05975 AA347404 AW069298 AI754351 AI754044  
AA188808 AA188879 AA565243 AL040655 AA456177 AI750722 AA045756 AA213580 C16936 AW578747 AW753731 H41632 N44761 R58560  
R61260 AA039902 N59721 AW992543 R68380 AA149686 T29017 H03739 BE383822 BE387105 BE408251 BE410425 H41660 AA247591  
BE389677 AI752233 AI568195 AA868004 AI424523 AW753720 AA852159 BE385803  
100559 2260\_1 NM\_000094 L02870 D13694 S51236 M96984 AW946290 M65158 AI285422 D29523 AL119886 AW630655 L06862 AI884355 AW168737 T29085  
AW797005 AW801340 AI355504 AW079048 AW801337 AI690455 AI972083 AW268565 W68588 AA587326 AA883498 AI033523 AW510356  
AW591998 H98463 AL043852 AI150055 AI566239 AI624803 AA844717 H40670 AA922334 AI864424 AW615094 AW451233 AI302203 F31221  
AI872170 W68589 AA904478 AI917631 AW014208 AW450759 AA847625 AI284033 AA848176 AA598507  
X00356 NM\_001741 M26095 X03662 M12667 X02330 X02330 AA716058 AW296074 X04861 AI695720 AA719597  
100576 9986\_1  
124357 genbank\_N22401 N22401  
101624 entrez\_M55998 M55998  
101625 entrez\_M57293 M57293  
135158 57963\_1 AL037551 AI804716 AW439811 AI569470 AA075299 AI738572 AI270388 AI816783 AW263026 AI633951 AI655285 AI990572 AI950425  
AW241533 AA916883 AA576693 AA160156 AA613783 AW078884 AI888282 AI275241 AI133467 AA164921

Tables 2A-8C were previously filed on November 9, 2001 in USSN 60/339,245 (18501-004100US)

Table 2A shows 504 genes down-regulated in lung tumors relative to normal lung and chronically diseased lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probe set obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

Pkey: Unique Eos probe set identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title  
 R1: 90th percentile of AI for normal lung samples divided by the 80th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.  
 R2: median of AI for normal lung samples divided by 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples.  
 R3: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinoma and squamous cell carcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.  
 R4: average of AI for normal lung samples divided by average AI for squamous cell carcinoma and adenocarcinoma lung tumors.  
 R5: median of AI for normal lung samples divided by the 90th percentile of AI for adenocarcinomas.  
 R6: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.  
 R7: average of AI for normal lung samples divided by the 90th percentile of AI for squamous cell carcinomas.  
 R8: median of AI for normal lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of AI for squamous cell carcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples.

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2	R3	R4	R5	R6	R7	R8
100095	Z97171	Hs.78454	myocilin; trabecular meshwork inducible	40.20							
100115	NM_002084	Hs.336920	glutathione peroxidase 3 (plasma)								3.46
100138	U83508	Hs.2463	angiopoietin 1			2.30					
100299	D49493	Hs.2171	growth differentiation factor 10		11.00						
100306	U86749	Hs.80598	transcription elongation factor A (SII);						3.06		
100447	NM_014767	Hs.74583	KIAA0275 gene product								3.16
100458	S74019	Hs.247979	Vpre-B	42.40							
100862	AA005247	Hs.285754	Hepatocyte Growth Factor Receptor						4.13		
100959	AA359129	Hs.118127	actin; alpha; cardiac muscle				125.60				
101032	BE205854	Hs.46039	phosphoglycerate mutase 2 (muscle)	36.40							
101081	AF047347	Hs.4880	amyloid beta (A4) precursor protein-bind				34.60				
101088	X70697	Hs.553	solute carrier family 6 (neurotransmitter)				193.20				
101125	AJ250562	Hs.82749	transmembrane 4 superfamily member 2						3.10		
101180	U11874	Hs.846	Interleukin 8 receptor; beta				54.86				
101308	L41390		*Homo sapiens core 2 beta-1,6-N-acetylgl	33.20							
101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do				36.40				
101345	NM_005795	Hs.152175	Calcitonin receptor-like			2.29					
101346	A1738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				70.55				
101397	M26380	Hs.180878	lipoprotein lipase								3.54
101414	NM_000066	Hs.38069	complement component 8; beta polypeptide							3.81	
101435	NM_001100	Hs.1288	actin; alpha 1; skeletal muscle				34.60				
101507	X16896	Hs.82112	interleukin 1 receptor; type I				37.60				
101530	M29874	Hs.1360	cytochrome P450; subfamily IIB (phenobar								4.25
101537	AA69059	Hs.184915	zinc finger protein; Y-linked			2.54					
101542	NM_000102	Hs.1363	cytochrome P450; subfamily XVII (steroid		5.50						
101545	BE245154	Hs.154210	EDG1; endothelial differentiation, sphin	39.40							
101554	BE207611	Hs.123078	thyroid stimulating hormone receptor		13.00						
101560	AW958272	Hs.83733	intercellular adhesion molecule 2, exon								3.38
101574	M34182	Hs.158029	protein kinase; cAMP-dependent; catalyti						4.37		
101605	M37984	Hs.118845	troponin C; slow								3.80
101621	BE391804	Hs.62661	guanylate binding protein 1; Interferon-	30.20							
101680	AA299330	Hs.1042	Sjogren syndrome antigen A1 (52kD; ribon							2.75	
101829	AW452398	Hs.129763	solute carrier family 8 (sodium/calcium						3.37		
101842	M93221	Hs.75182	mannose receptor; C type 1				38.20				
101961	AW004056	Hs.168357	*Hs-TBX2=T-box gene (T-box region) [huma			2.32					
101994	T92248	Hs.2240	uteroglobin								6.85
102020	AU077315	Hs.154970	transcription factor CP2			2.45					
102091	BE280901	Hs.83155	aldehyde dehydrogenase 7								6.75
102112	AW025430	Hs.156591	forkhead box F1	54.60							
102190	AA723157	Hs.73769	folate receptor 1 (adult)								3.98
102202	NM_000507	Hs.574	fructose-bisphosphatase 1								3.62
102241	NM_007351	Hs.268107	Multimerin			2.32					
102310	U33839		Accession not listed in Genbank		7.00						
102397	U41898		*Human sodium cotransporter RKST1 mRNA,	29.40							
102571	U60115	Hs.239069	*Homo sapiens skeletal muscle LIM-protei								3.75
102620	AA976427	Hs.121513	Human clone Y2-6 mRNA from chromosome X						3.07		
102636	U67092		*Human ataxia-telangiectasia locus prote			2.40					
102667	U70867	Hs.83974	solute carrier family 21 (prostaglandin			3.15					
102675	U72512	Hs.7771	*Human B-cell receptor associated protei						3.56		
102698	M18667	Hs.1867	progastrin (pepsinogen C)								4.51
102727	U79251	Hs.99902	opioid-binding protein/cell adhesion mol					12.00			
102852	V00571	Hs.75294	corticotropin releasing hormone	37.40							
103026	X54162	Hs.79386	thyroid and eye muscle autoantigen D1 (6					13.00			
103028	X54380	Hs.74094	pregnancy-zone protein	28.80							
103098	M86361		Human mRNA for T cell receptor; clone IG					10.00			
103117	X63578	Hs.295449	parvalbumin		6.00						
103241	X76223		H.sapiens MAL gene exon 4			2.47					
103280	U84722	Hs.76206	Cadherin 5, VE-cadherin (vascular epithe			2.69					
103360	Y16791	Hs.73082	keratin; hair; acidic; 5							2.16	



## PCT/US02/12476

5	103496	Y09267	Hs.132821	flavin containing monooxygenase 2					5.97
	103508	Y10141		"H.sapiens DAT1 gene, partial, VNTR"			3.27		
	103561	NM_001843	Hs.143434	contactin 1	2.40				
	103569	NM_005512	Hs.151641	glycoprotein A repetitions predominant	2.99				
	103575	Z26256		"H.sapiens isoform 1 gene for L-type cal			4.18		
	103627	Z48513		H.sapiens XG mRNA (clone PEP6)			3.44		
	103767	BE244667	Hs.296155	CGI-100 protein				2.25	
	103850	AA187101	Hs.213194	Hypothetical protein MGC10895; sim to SR		46.55			
10	104078	AA402801	Hs.303276	ESTs			3.05		
	104326	AW732858	Hs.143067	ESTs			3.54		
	104352	BE219898	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl			3.16		
	104398	AI423930	Hs.36790	ESTs; Weakly similar to putative p150 [H	64.80				
	104473	AI904823	Hs.31297	ESTs					3.38
15	104493	AW960427	Hs.79059	ESTs; Moderately similar to TGF-BETA REC		2.47			
	104495	AW975687	Hs.292979	ESTs	28.60				
	104595	AI799603	Hs.271568	ESTs			3.42		
	104597	AI364504	Hs.53957	ESTs; Weakly similar to Silt-1 protein [		6.00			
	104659	AW969769	Hs.105201	ESTs	34.00				
20	104686	AA010539	Hs.18912	ESTs		11.00			
	104691	U29690	Hs.37744	ESTs; Beta-1-adrenergic receptor	56.80				
	104764	AI039243	Hs.278585	ESTs			60.40		
	104776	AA026349		ESTs	34.20				
	104825	AA035613	Hs.141883	ESTs		3.03			
25	104855	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C	41.20				
	104942	NM_016348	Hs.10235	ESTs					3.27
	104989	R65998	Hs.285243	ESTs			40.00		
	105062	AW954355	Hs.36529	ESTs					3.20
	105101	H63202	Hs.38163	ESTs	34.20				
30	105173	U54617	Hs.8364	ESTs					4.17
	105194	R08780	Hs.19800	ESTs		16.00			
	105226	R58958	Hs.28608	ESTs			2.34		
	105256	AA430650	Hs.16529	transmembrane 4 superfamily member (tet			2.72		
	105394	BE245812	Hs.8941	ESTs			2.61		
35	105647	Y09306	Hs.30148	homeodomain-interacting protein kinase 3	33.60				3.59
	105789	AF106941	Hs.18142	arrestin; beta 2					
	105817	AA397825		synaptopodin				4.46	
	105847	AW964490	Hs.32241	ESTs					
	105894	AI904740	Hs.25691	calcitonin receptor-like receptor activi		3.43			
40	105999	BE268786	Hs.21543	ESTs		7.00			
	106075	AA045290	Hs.25830	ESTs			42.60		
	106178	AL049935	Hs.301763	KIA0554 protein	34.80				
	106381	AB040916	Hs.24106	ESTs			12.00		
	106467	AA450040	Hs.154162	ADP-ribosylation factor-like 2				3.69	
45	106636	AA329648	Hs.23804	ESTs			96.40		
	106569	R20909	Hs.300741	sorcin			47.20		
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr			220.40		
	106842	AF124251	Hs.26054	novel SH2-containing protein 3		2.55			
	106844	AA480505	Hs.158213	sperm associated antigen 6	39.20				
50	106870	AI983730	Hs.26530	serum deprivation response (phosphatidy)		2.28			
	106943	AW888222	Hs.5973	ESTs					4.28
	106954	AF128847	Hs.204038	ESTs					4.32
	107106	AA862496	Hs.28482	ESTs			10.45		
	107163	AF233588	Hs.27018	ESTs		2.57			
55	107201	D20378	Hs.30731	EST				3.84	
	107238	D59362	Hs.330777	EST		8.00			
	107376	U90545	Hs.327179	solute carrier family 17 (sodium phospho		10.67			
	107530	Y13622	Hs.85087	latent transforming growth factor beta b					
	107688	AW082221	Hs.60536	ESTs			34.60		
60	107706	AA015579	Hs.29276	ESTs	28.40				
	107723	AA015967		EST				3.29	
	107727	AA149707	Hs.173091	DKFZP434K151 protein			80.80		
	107750	AA017291	Hs.60781	ESTs			51.40		
	107751	AA017301	Hs.235390	ESTs				3.14	
65	107873	AK000520	Hs.143811	ESTs		9.00			
	107899	BE019261	Hs.83869	ESTs; Weakly similar to I III ALU SUBFAM I				3.65	
	107994	AA036811	Hs.48469	ESTs			44.60		
	107997	AL049176	Hs.82223	Human DNA sequence from clone 141H5 on c			32.00		
	108041	AW204712	Hs.61957	ESTs			30.80		
70	108048	AI797341	Hs.165195	ESTs					4.75
	108338	AA070773		*zm53g11.s1 Stratagene fibroblast (#9372		2.33			
	108434	AA078899		*zm94b1.s1 Stratagene colon HT29 (#93722					2.92
	108447	AA079126		*zm92a11.s1 Stratagene ovarian cancer (#				3.06	
	108480	AL133092	Hs.68055	ESTs			34.00		
75	108499	AA083103		*zn1b12.s1 Stratagene hNT neuron (#93723					3.36
	108535	R13949	Hs.226440	Homo sapiens clone 24881 mRNA sequence			19.00		
	108550	AA084867		*zn11f6.s1 Stratagene hNT neuron (#93723			12.00		
	108604	AA934589	Hs.49696	ESTs		2.33			
	108625	AW972330	Hs.283022	ESTs					5.82
80	108629	AA102425		*zn24c6.s1 Stratagene neuroepithelium NT				3.42	
	108655	AA09960		*zm65c6.s1 Stratagene fibroblast (#93721		7.00			
	108756	AA127221	Hs.117037	Homo sapiens mRNA; cDNA DKFZp564N1164 (f		6.05			
	108864	AI733852	Hs.199597	ESTs	28.80				
	108895	AL138272	Hs.62713	ESTs	32.80				
	108921	AI568801	Hs.71721	ESTs			57.80		
85	108957	AA142989	Hs.71730	ESTs	28.80				

## WO 02/086443

## PCT/US02/12476

	109001	AI056548	Hs.72116	ESTs, Moderately similar to hedgehog-int	2.57				
	109003	AA147497	Hs.71825	ESTs				2.11	
	109004	AA156235	Hs.139077	EST	5.60				
5	109065	AA161125	Hs.252739	EST			10.00		
	109250	H83784	Hs.62113	ESTs; Weakly similar to PHOSPHATIDYLETHA				3.44	
	109490	AA233416	Hs.139202	ESTs				2.92	
	109510	AJ798863	Hs.87191	ESTs	2.40				
	109578	F02208	Hs.27214	ESTs	10.00				
10	109601	F02695	Hs.311662	EST			40.80		
	109613	H47315	Hs.27519	ESTs			54.40		
	109650	R31770	Hs.23540	ESTs	31.20				
	109682	H18017	Hs.22869	ESTs	8.40				
	109724	D59899	Hs.127842	ESTs			29.40		
15	109782	AB020644	Hs.14945	long fatty acyl-CoA synthetase 2 gene			8.00		
	109833	R79854	Hs.29889	ESTs	10.00				
	109837	H00656	Hs.29792	ESTs	6.49				
	109977	T64183	Hs.282982	ESTs				2.75	
	109984	AJ796320	Hs.10299	ESTs			107.00		
20	110146	H41324	Hs.31581	ESTs; Moderately similar to SYNTAXIN 1B				2.22	
	110271	H28985	Hs.31330	ESTs				3.48	
	110280	AW674263	Hs.32468	ESTs	44.20				
	110420	R93141	Hs.184261	ESTs			32.00		
	110578	T62507	Hs.11038	ESTs	28.40				
25	110634	R98905	Hs.35992	ESTs			20.00		
	110726	AW961818	Hs.24379	potassium voltage-gated channel; shaker-				4.15	
	110837	H03109	Hs.108920	ESTs; Weakly similar to semaphorin F [H.			56.80		
	110875	N35070	Hs.26401	tumor necrosis factor (ligand) superfamily	3.13				
	110894	R92356	Hs.66881	ESTs; Moderately similar to cytoplasmic	5.33				
30	110971	AJ760098	Hs.21411	ESTs			44.60		
	111023	AV655386	Hs.7645	ESTs	32.40				
	111057	T79639	Hs.14629	ESTs			17.14		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f				4.58	
	111330	BE247767	Hs.18166	KIAA0870 protein				3.42	
35	111374	BE250726	Hs.283724	ESTs; Moderately similar to HYA22 [H.sap				3.91	
	111442	AW449573	Hs.181003	ESTs			33.20		
	111737	H04607	Hs.9218	ESTs			53.00		
	111747	AJ741471	Hs.23666	ESTs	46.20				
	111807	R33508	Hs.18827	ESTs	16.00				
40	111862	R37472	Hs.21559	EST				3.91	
	112045	AI372588	Hs.8022	TU3A protein				2.74	
	112057	R43713	Hs.22945	EST				4.92	
	112214	AW148652	Hs.167398	ESTs			13.00		
	112263	R52393	Hs.25917	ESTs		2.43			
45	112314	AW206093	Hs.748	ESTs	9.00				
	112324	R55965	Hs.26479	limbic system-associated membrane protein			14.00		
	112362	AW300887	Hs.26638	ESTs; Weakly similar to CD20 receptor [H		2.49			
	112380	H63010	Hs.5740	ESTs		2.34			
	112425	AA324998	Hs.321677	ESTs; Weakly similar to III ALU SUBFAMI	8.00				
50	112473	R65993	Hs.279798	pregnancy specific beta-1-glycoprotein 9				4.53	
	112492	N51620	Hs.28694	ESTs			29.80		
	112541	AF038392	Hs.116674	ESTs				3.62	
	112620	R80652	Hs.29040	ESTs	2.37				
	112623	AW373104	Hs.25094	ESTs	2.26				
55	112867	T03254	Hs.167393	ESTs			12.00		
	112894	T08188	Hs.3770	ESTs	6.50				
	112854	AA928953	Hs.6555	ESTs	7.00				
	113029	AW081710	Hs.7369	ESTs; Weakly similar to III ALU SUBFAMI				4.39	
	113086	AA346839	Hs.209100	DKFZP434C171 protein				4.47	
60	113140	T50405	Hs.175967	ESTs			10.00		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	14.00				
	113257	AI821378	Hs.159367	ESTs				3.72	
	113394	T61473	Hs.177894	ESTs				3.60	
	113437	T85349	Hs.15923	EST	35.00				
65	113454	AJ022166	Hs.16188	ESTs	6.00				
	113502	T89130		ESTs	39.60				
	113552	AI654223	Hs.16026	ESTs					3.88
	113645	T95358	Hs.333181	ESTs				2.58	
	113691	T96935	Hs.17932	EST			38.20		
70	113706	AA004693	Hs.269192	ESTs				3.09	
	113883	U89281	Hs.11958	oxidative 3 alpha hydroxysteroid dehydro		2.31			
	113924	BE178285	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	30.40				
	114035	W92798	Hs.269181	ESTs			13.00		
	114058	AK002016	Hs.114727	ESTs					5.00
75	114084	AA708035	Hs.12248	ESTs			40.60		
	114121	H05785	Hs.25425	ESTs		2.31			
	114124	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)	7.00				
	114275	AW515443	Hs.306117	interleukin 13 receptor; alpha 1	6.00				
	114297	AA149707	Hs.173091	DKFZP434K151 protein			48.80		
80	114427	AA017176	Hs.33532	ESTs; Highly similar to Miz-1 protein [H				3.45	
	114449	AA020736		*ze63b11.s1 Soares retina N2b4HR Homo sa			10.00		
	114452	AI369275	Hs.243010	ESTs, Moderately similar to RTCO_HUMAN G	14.00				
	114609	AA079505		*zm57a5.s1 Stratagene colon HT29 (#93722				3.13	
	114648	AA101056		*zn25b3.s1 Stratagene neuroepithelium NT			35.40		
85	114731	BE094291	Hs.155651	Homo sapiens HNF-3beta mRNA for hepatocy				3.42	
	114762	AA146979	Hs.288464	ESTs	33.00				



## WO 02/086443

## PCT/US02/12476

	114776	AA151719	Hs.95834	ESTs	34.40				
	115009	AA251561	Hs.48689	ESTs	30.20				
	115272	AW015947		ESTs; Weakly similar to hypothetical L1	32.60				
5	115279	AW964897	Hs.290825	ESTs		6.00			
	115302	AL109719	Hs.47578	ESTs			12.00		
	115365	AW976252	Hs.268391	ESTs				3.32	
	115559	AL079707	Hs.207443	ESTs			48.00		
	115566	AI142336	Hs.43977	ESTs			56.20		
10	115683	AF255910	Hs.54650	ESTs; Weakly similar to (define not ava	31.40				
	115744	AA418538	Hs.43945	ESTs; Highly similar to dJ1178H5.3 [Hsa			33.60		
	115819	AA486620	Hs.41135	Endomucin 2			74.40		
	115949	AI478427	Hs.43125	ESTs		3.18			
	115965	AA001732	Hs.173233	ESTs			388.80		
15	116035	AA621405	Hs.184664	ESTs			33.20		
	116049	AA454033	Hs.41644	ESTs			45.80		
	116081	AI190071	Hs.55278	ESTs				3.57	
	116082	AB029496	Hs.59729	ESTs		3.06			
	116213	AA292105	Hs.326740	leucine rich repeat (in FLI) Interactin	50.60				
20	116228	AI767947	Hs.50341	ESTs; Weakly similar to tuftelin [M.musc		3.85			
	116250	N76712	Hs.44829	ESTs		6.00			
	116419	AI613480	Hs.47152	ESTs; Weakly similar to testicular tek1			30.00		
	116617	D60761	Hs.45220	EST		2.27			
	116784	AB007979	Hs.301281	tenascin R (restriclin; janusin)	47.20				
25	116835	N39230	Hs.38218	ESTs			41.20		
	116970	AB023179	Hs.9059	KIAA0962 protein			11.00		
	117023	AW070211	Hs.102415	ESTs			91.00		
	117027	AW085208	Hs.130093	ESTs	49.40				
	117036	H88908	Hs.41192	EST			32.60		
30	117110	AA160079	Hs.172932	ESTs	8.67				
	117209	W03011	Hs.306881	ESTs			30.60		
	117325	N23599	Hs.43396	ESTs				9.29	
	117454	N29589	Hs.44055	ESTs					3.19
	117475	N30205	Hs.93740	ESTs	44.00				
35	117543	BE219453	Hs.42722	ESTs		16.00			
	117567	AW444761	Hs.44565	ESTs			12.00		
	117570	N48649	Hs.44583	ESTs			11.00		
	117600	N34963	Hs.44676	EST				3.74	
	117730	N45513	Hs.46608	ESTs		6.00			
40	117791	N48325	Hs.93956	EST		9.00			
	117929	N51075	Hs.47191	ESTs			29.20		
	117990	AA446167	Hs.47385	ESTs		8.00			
	118224	N62275	Hs.48503	EST	31.40				
	118244	N62516	Hs.48556	ESTs	32.80				
45	118357	AL109667	Hs.124154	Homo sapiens mRNA full length insert cDN		2.40			
	118446	N66361	Hs.269121	ESTs		2.28			
	118447	N66399	Hs.49193	EST	30.80				
	118530	N67900	Hs.118446	ESTs				3.10	
	118549	N68163	Hs.322954	EST				3.41	
50	118823	W03754	Hs.50813	ESTs; Weakly similar to long chain fatty		3.94			
	118862	W17065	Hs.54522	ESTs				3.58	
	118935	AI979247	Hs.247043	KIAA0525 protein			33.00		
	118944	AI734233	Hs.226142	ESTs; Weakly similar to III ALU SUBFAM				11.43	
	118995	N94591	Hs.323056	ESTs		14.00			
55	119073	BE245360	Hs.279477	ERG-2/ERG-1; V-ets avian erythroblastosi			52.60		
	119268	T16335	Hs.65325	EST	31.40				
	119514	W37937		Accession not listed in Genbank				3.50	
	119824	W74536	Hs.184	advanced glycosylation end product-sped		2.75			
60	119831	AL117664	Hs.58419	DKFZP586L2024 protein					3.21
	119861	W78816	Hs.49943	ESTs; Moderately similar to III ALU SUB			33.80		
	119889	W84346	Hs.58671	ESTs			30.03		
	119921	W86192	Hs.58815	ESTs	29.00				
	120082	H80286	Hs.40111	ESTs				3.80	
	120094	AA811339	Hs.124049	ESTs		6.00			
65	120132	W57554	Hs.125019	Human lymphoid nuclear protein (LAF-4)			36.60		
	120378	AA223249	Hs.285728	ESTs		12.00			
	120404	AB023230	Hs.96427	KIAA1013 protein	39.40				
	120504	AA256837		ESTs			8.00		
	120512	N55761	Hs.194718	ESTs	33.00				
70	120667	AA287740	Hs.78335	microtubule-associated protein; RP/EB fa				4.18	
	120777	AA287702	Hs.10031	KIAA0955 protein			46.60		
	121082	AA398722		ESTs			39.00		
	121191	AA400205	Hs.104447	ESTs	41.60				
	121248	AA400914	Hs.97827	EST				5.08	
75	121363	AI287280	Hs.97933	ESTs			12.00		
	121366	AI743515		ESTs			20.00		
	121483	AI680332	Hs.25274	ESTs; Moderately similar to putative sev				3.32	
	121518	AA412155		ESTs			30.20		
	121545	AA412442	Hs.98132	ESTs		2.29			
80	121622	AA416931	Hs.126065	ESTs		9.00			
	121665	AA416556	Hs.98234	ESTs			34.80		
	121709	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	34.80				
	121730	AI140683	Hs.98328	ESTs					
	121740	AA421138	Hs.98334	EST		7.00			
85	121772	AI590770	Hs.110347	Homo sapiens mRNA for alpha Integrin bin	36.20				
	121821	AL040235	Hs.3346	ESTs				3.61	

## WO 02/086443

## PCT/US02/12476

	121835	AB033030	Hs.300670	ESTs	2.34				
	121841	AA427794	Hs.104864	ESTs	2.61				
	121885	AA934883	Hs.98467	ESTs				2.25	
	121888	AA426429	Hs.98463	ESTs				2.92	
5	121938	AA428659	Hs.98610	ESTs			46.80		
	121950	AA429515		EST			31.40		
	122030	AA431310	Hs.98724	ESTs	34.40				
	122054	AA431725	Hs.98746	EST				3.58	
10	122211	AA300900	Hs.98849	ESTs; Moderately similar to bithoraxoid-	49.40				
	122233	AA435455	Hs.98872	EST	29.80				
	122247	AA436676	Hs.98890	EST			39.80		
	122253	AA436703	Hs.104938	ESTs; Weakly similar to hypothetical pro	9.00				
	122266	AA436840	Hs.98907	EST				3.60	
	122285	AA436981	Hs.121602	EST				3.14	
15	122409	AA446830	Hs.99081	ESTs	30.80				
	122485	AA524547	Hs.160318	phospholemman		2.65			
	122697	AA420683	Hs.98321	Homo sapiens cDNA FLJ14103 fis, clone MA	15.00				
	122772	AW117452	Hs.99489	ESTs	6.67				
20	122831	AI857570	Hs.5120	ESTs				3.37	
	122913	AI638774	Hs.105328	ESTs			32.20		
	123049	BE047680	Hs.211869	ESTs			41.80		
	123076	AI345569	Hs.190046	ESTs	35.80				
	123136	AW451999	Hs.194024	ESTs				2.58	
25	123309	N52937	Hs.102679	ESTs			19.00		
	123455	AA353113	Hs.112497	ESTs			82.80		
	123691	AA609579	Hs.112724	ESTs				3.95	
	123756	AA609971	Hs.112795	EST	35.40				
	123802	AA620448		Homo sapiens clone 24760 mRNA sequence	58.00				
30	123837	AI807243	Hs.112893	ESTs			32.40		
	123844	AA938905	Hs.120017	olfactory receptor, family 7; subfamily		2.63			
	123936	NM_004673	Hs.241519	ESTs	29.00				
	123987	C21171	Hs.95497	ESTs; Weakly similar to GLUCOSE TRANSPOR			70.60		
	124013	AI521936	Hs.107149	ESTs; Weakly similar to PTB-ASSOCIATED S	28.40				
35	124160	R40290	Hs.124685	ESTs			13.00		
	124205	H77570	Hs.108135	ESTs				4.74	
	124226	AA618527	Hs.190266	ESTs		2.35			
	124246	H87680	Hs.270862	ESTs			29.40		
	124348	AI96320	Hs.10299	ESTs	17.00				
40	124358	AW070211	Hs.102415	gamma35g11.s1 Morton Fetal Cochlea Homo sa		3.07			
	124409	AI814166	Hs.107197	ESTs				3.14	
	124442	AW663632	Hs.285625	TATA box binding protein (TBP)-associate		2.48			
	124468	N51413	Hs.109284	ESTs			30.80		
	124479	AB011130	Hs.127436	calcium channel; voltage-dependent; alph				6.03	
45	124519	AI670056	Hs.137274	ESTs; Weakly similar to SPLICEOSOME ASSO		2.50			
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	59.20				
	124866	AI768289	Hs.304389	ESTs	8.00				
	124874	BE550182	Hs.127826	ESTs			37.60		
	125097	AW576389	Hs.335774	ESTs			10.00		
50	125179	AW206468	Hs.103118	ESTs				3.12	
	125200	AW836591	Hs.103166	ESTs				2.79	
	125299	T32982	Hs.102720	ESTs			34.20		
	125400	AL110151	Hs.128797	DKFZP586D0824 protein	29.00				
	125810	H00083		aryl hydrocarbon receptor-interacting pr	32.20				
55	126176	BE242256	Hs.2441	KIAA0022 gene product		12.00			
	126303	D78841		HUM525A05B Human placenta polyA+ (TFu)			33.60		
	126403	AW629054	Hs.126976	ESTs; Weakly similar to metalloprotease/	35.80				
	126507	AL040137	Hs.23964	ESTs; Weakly similar to HC1 ORF [M.muscu			29.80		
	126773	AA648284	Hs.187584	ESTs	39.60				
60	127307	AW962712	Hs.126712	ESTs; Weakly similar to pIL2 hypotheticala	28.80				
	127462	AA760776	Hs.293977	aa59b04.s1 NCL_CGAP_GCB1 Homo sapiens c			34.40		
	127486	AW002846	Hs.105468	ESTs	9.00				
	127572	AA594027	Hs.191788	ESTs		2.36			
	127609	X80031	Hs.530	ESTs			29.40		
65	127832	AW976035	Hs.292396	ESTs			37.20		
	127898	AA774725	Hs.126970	ESTs				4.42	
	128073	AW340720	Hs.126983	ESTs			38.40		
	128101	AA905730	Hs.128254	ESTs	7.33				
	128149	NM_012214	Hs.177576	mannosyl (alpha-1,3)-glycoprotein beta-				2.58	
70	128212	W27411	Hs.336920	glutathione peroxidase 3 (plasma)		3.09			
	128333	W68800	Hs.12126	ESTs; Weakly similar to LR8 (H.sapiens)			34.40		
	128364	N76462	Hs.269152	ESTs; Weakly similar to ZINC FINGER PROT		10.00			
	128426	AI265784	Hs.145197	ESTs				4.31	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel; s	31.20				
75	128634	AA464918		ESTs; Moderately similar to IIII ALU SUB			41.60		
	128687	AW271273	Hs.23767	ESTs			87.00		
	128726	AI311238	Hs.104476	ESTs				4.02	
	128773	NM_004131	Hs.1051	granzyme B (granzyme 2; cytotoxic T-lymp			9.00		
	128833	W26667	Hs.184581	ESTs				3.76	
80	128870	H39537	Hs.75309	eukaryotic translation elongation factor		2.66			
	128878	R25513	Hs.10583	ESTs				3.10	
	128885	AF134803	Hs.180141	cofilin 2 (muscle)			11.00		
	128998	W04245	Hs.107761	ESTs; Weakly similar to PUTATIVE RHO/RAC				3.21	
	129000	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII Inhi				3.68	
	129038	AW156903	Hs.108124	ribosomal protein L41				3.17	
85	129098	AW580945	Hs.330466	ESTs	34.60				

WO 02/086443

PCT/US02/12476

	129210	AL039940	Hs.202949	KIAA1102 protein				4.09
	129240	AA361258	Hs.237868	interleukin 7 receptor	2.29			
	129262	BE222198	Hs.109843	ESTs		3.30		
5	129301	AF182277	Hs.330780	Human cytochrome P450-11B (h11B3) mRNA;				4.05
	129331	AW167669	Hs.279772	ESTs; Highly similar to CGI-38 protein [				4.09
	129381	AW245805	Hs.110903	claudin 5 (transmembrane protein deleted	2.93			
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		160.80		
	129595	U09550	Hs.11154	oviductal glycoprotein 1; 120kD		10.00		
10	129613	AW978517	Hs.172847	ESTs; Weakly similar to collagen alpha 1			3.40	
	129782	AW016932	Hs.104105	EST	9.00			
	129950	F07783	Hs.1369	decay accelerating factor for complement		87.80		
	129958	R27496	Hs.1378	annexin A3		44.60		
	129959	AL036554	Hs.274463	defensin; alpha 1; myeloid-related seque	2.72			
15	130160	AA305688	Hs.267695	UDP-GalbetaGlcNAc beta 1,3-galactosyltr		42.20		
	130259	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	2.54			
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic; Dr		51.60		
	130312	AF056195	Hs.15430	DKFZP586G1219 protein			3.16	
	130436	NM_001928	Hs.155597	D component of complement (adipsin)				4.11
20	130523	AA999702	Hs.214507	ESTs			4.77	
	130799	AB028945	Hs.12696	ESTs	6.00		3.54	
	130885	NM_005883	Hs.20912	adenomatous polyposis coli like				3.50
	131002	AL050295	Hs.22039	KIAA0758 protein				
	131012	AL039940	Hs.202949	KIAA1102 protein	20.00			
25	131031	NM_001650	Hs.288650	aquaporin 4	41.20			
	131061	N64328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.		31.40		
	131066	AW169287	Hs.22588	ESTs		29.60		
	131082	AB091121	Hs.246218	ESTs; Weakly similar to zinc finger prot		9.00		
	131087	AF147709	Hs.22824	ESTs; Weakly similar to p160 myb-binding				3.86
30	131161	AF033382	Hs.23735	potassium voltage-gated channel; subfam			3.14	
	131179	AA171388	Hs.184482	DKFZP586D0624 protein			3.80	
	131182	AB24144	Hs.23912	ESTs				3.67
	131205	NM_003102	Hs.2420	superoxide dismutase 3; extracellular	2.98			
	131277	AA131466	Hs.23767	ESTs	3.15			
35	131281	AA251716	Hs.25227	ESTs		32.20		
	131282	X03350	Hs.4	alcohol dehydrogenase 3 (class I); gamma				3.44
	131285	AB57943	Hs.25274	ESTs; Moderately similar to putative sev			6.40	
	131355	R52804	Hs.25956	DKFZP564D206 protein	8.00			
	131391	AW085781	Hs.26270	ESTs	10.00			
40	131461	AA992841	Hs.27263	butyrate response factor 2 (EGF-response	28.80			
	131487	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f				4.03
	131517	AB037789	Hs.263395	ESTs; Highly similar to semaphorin 11a [	39.00			
	131545	AL137432	Hs.28564	ESTs		11.00		
	131583	AK000383	Hs.323092	ESTs; Weakly similar to dual specificity		10.00		
45	131647	AA359615	Hs.30089	ESTs	2.47			
	131675	H15205	Hs.30509	ESTs			3.06	
	131676	AB126821	Hs.30514	ESTs	45.80			
	131708	S60415	Hs.30941	calcium channel; voltage-dependent; beta	2.28			
	131717	X94630	Hs.3107	CD97 antigen				3.78
50	131766	AA443966	Hs.31595	ESTs		40.60		
	131762	AA744902	Hs.107767	ESTs; Moderately similar to CaM-KII inh				3.67
	131821	AA017247	Hs.164577	ESTs	2.87			
	131839	AB014533	Hs.33010	KIAA0633 protein			3.48	
	131861	AL056858	Hs.184245	KIAA0929 protein Mx2 interacting nuclea	54.00			
55	132015	AI418006	Hs.3731	ESTs		49.20		
	132070	BE622641	Hs.38489	ESTs		34.80		
	132242	AA332697	Hs.42721	ESTs	2.68			
	132334	AW080704	Hs.45033	lacritinal proline rich protein	4.66			
	132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Cri-du-chat regi	34.20			
60	132490	NM_001290	Hs.4980	UIM binding domain 2		2.66		
	132533	AB22988	Hs.172510	ESTs	13.00			
	132598	X80031	Hs.530	collagen; type IV; alpha 3 (Goodpasture		30.60		
	132619	H28855	Hs.53447	ESTs; Moderately similar to kinesin ligh			4.02	
	132652	N41739	Hs.61260	ESTs			3.18	
65	132726	N52298	Hs.55608	ESTs; Weakly similar to cDNA EST yk484g1		11.43		
	133028	R51604	Hs.300842	ESTs	2.37			
	133071	BE384932	Hs.64313	ESTs	2.27			
	133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein	2.63			
	133129	AA428580	Hs.65551	ESTs				5.49
70	133147	AA026533	Hs.66	interleukin 1 receptor-like 1	6.20			
	133151	NM_014051	Hs.94896	ESTs			3.69	
	133213	AA903424	Hs.6786	ESTs		31.40		
	133276	AW978439	Hs.69504	ESTs		9.00		
	133377	AI131245	Hs.7239	SEC24 (S. cerevisiae) related gene famil	41.20			
75	133407	AF017987	Hs.7305	secreted frizzled-related protein 1	50.20			
	133535	AL134030	Hs.284180	protocadherin 2 (cadherin-like 2)			3.72	
	133537	U41518	Hs.74602	aquaporin 1 (channel-forming integral pr				3.35
	133656	BE149455	Hs.75415	Accession not listed in Genbank	2.65			
	133689	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)		90.80		
80	133779	T68486	Hs.222566	ESTs			3.05	
	133978	AF035718	Hs.78051	transcription factor 21	2.92			
	133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec				3.45
	134000	AW175787	Hs.334841	selenium binding protein 1				4.05
	134111	AI372688	Hs.8022	TU3A protein	4.49			
85	134185	AA285136	Hs.301914	Homo sapiens mRNA; cDNA DKFZp586K1220 (f			3.27	
	134204	AB73257	Hs.7994	ESTs; Weakly similar to CGI-69 protein [		40.80		

## WO 02/086443

## PCT/US02/12476

	134641	AI092634	Hs.156114	protein tyrosine phosphatase; non-recept					
	134677	AA251363	Hs.177711	ESTs		32.20		3.76	
	134745	NM_000685	Hs.89472	angiotensin receptor 1B	15.00				
5	134749	T28499	Hs.89485	carbonic anhydrase IV		3.05			
	134786	T28618	Hs.89640	angiotensin 1 receptor; TEK tyrosine ki			57.80		
	134825	U33749	Hs.197764	thyroid transcription factor 1					3.73
	134978	AI829008	Hs.333383	ficollin (collagen/fibrinogen domain-cont		2.52			
	135010	N50465	Hs.92927	ESTs			31.60		
10	135053	AW796190	Hs.93678	ESTs				3.21	
	135081	AF069517	Hs.173993	RNA binding motif protein 6	28.80				
	135091	AA493650	Hs.94367	ESTs					4.24
	135135	AA775910	Hs.95011	syntrophin; beta 1 (dystrophin-associate		8.00			
	135203	C15737	Hs.269386	ESTs				4.31	
15	135236	AI636208	Hs.96901	ESTs					
	135266	R41179	Hs.97393	Human mRNA for KIAA0328 gene; partial cd	43.00			6.42	
	135346	NM_000928	Hs.992	phospholipase A2; group IB (pancreas)		3.82			
	135378	AW951818	Hs.24379	potassium voltage-gated channel; shaker-		4.15			
	135387	NM_001972	Hs.99863	elastase 2; neutrophil	37.20				
20	135388	WZ7965	Hs.99865	EST	38.80				
	135402	L12398	Hs.99922	dopamine receptor D4				4.21	

TABLE 2B shows the accession numbers for those primekeys lacking unigenelD's for Table 2A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
108447	43452_7	AA079126
108550	120073_1	AA084857 AA084996
108655	127522_1	AA099960 AA113013
102397	44371_1	U41898
126303	1525933_1	D78841 D78880
125810	1554054_1	H00083 R81062
103627	2615_2	Z48513 Z48512
121366	280401_1	AI743515 AA405617 AW276706
114609	116777_1	AA079505 AA079537
115272	172113_1	AW015947 AA211890 AA279425
108338	112186_1	AA070773 AA070774
108434	114012_1	AA078899 AA078782 AA075788
123802	genbank_AA620448	AA620448
102310	NOT_FOUND_entrez_U33839	U33839
102636	entrez_U67092 U67092	
104776	genbank_AA026349	AA026349
120504	genbank_AA256837	AA256837
113502	genbank_T89130T89130	
108499	genbank_AA083103	AA083103
101308	entrez_L41390 L41390	
108629	genbank_AA102425	AA102425
103098	Z21_215	M86361 Z26593 X02850 D13070 AE000659 M17649 M87869 M87871 X61077 M16286 AF018169 X61079 S59351 X60142 AF043169
103241	entrez_X76223 X76223	
103508	entrez_Y10141 Y10141	
103575	entrez_Z26256 Z26256	
119514	NOT_FOUND_entrez_W37937 W37937	
121082	genbank_AA398722	AA398722
128634	AA464918_et AA464918	
105817	genbank_AA397825	AA397825
121518	genbank_AA412155	AA412155
114449	genbank_AA020736	AA020736
114648	genbank_AA101056	AA101056
121950	genbank_AA429515	AA429515
107723	genbank_AA015957	AA015957

Table 3A shows 452 genes up-regulated in chronically diseased lung relative to normal lung. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

- 5 Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title  
 10 R1: 80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of AI for normal lung samples.  
 R2: 80th percentile of AI for chronically diseased lung samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas  
 R3: 70th percentile of AI for chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by the 90th percentile of normal lung samples, squamous cell carcinomas and adenocarcinomas minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples

15

Pkey	ExAccn	UnigenelD	Unigene Title	R1	R2	R3
135423	U50531	Hs.138751	Human BRCA2 region, mRNA sequence CG030	12.40		
135378	AW561818	Hs.24379	MUM2 protein			2.13
135346	NM_000928	Hs.992	phospholipase A2, group IB (pancreas)			
135235	AW298244	Hs.293507	ESTs	12.40		
135057	U90268	Hs.93810	cerebral cavernous malformations 1	11.67		
134951	BE305081	Hs.169358	hypothetical protein		8.00	
134799	M36821	Hs.89690	GRO3 oncogene		8.20	
134786	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous)			
134772	NM_000829	Hs.163697	glutamate receptor, ionotropic, AMPA 4	29.80		
134752	BE246762	Hs.89499	arachidonate 5-lipoxygenase			1.93
134749	T28499	Hs.89485	carbonic anhydrase IV			2.07
134696	BE326276	Hs.88861	ESTs			
134636	NM_005582	Hs.87205	lymphocyte antigen 64 (mouse) homolog, r	13.60		
134627	A018768	Hs.12482	glyceronephosphate O-acyltransferase			1.92
134622	AW575159	Hs.293097	ESTs, Weakly similar to A55380 facicogeni			1.92
134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	13.20		
134561	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h			1.78
134468	NM_001772	Hs.83731	CD33 antigen (gp67)		6.20	
134417	NM_006416	Hs.82921	solute carrier family 35 (CMP-eatic aci			
134343	D50683	Hs.82028	transforming growth factor, beta recepto			
134323	BE170651	Hs.8700	deleted in liver cancer 1			
134300	NM_001430	Hs.8136	endothelial PAS domain protein 1			
134299	AW580939	Hs.97199	complement component C1q receptor			
134253	X52075	Hs.80738	slalophorin (gpl.115, leukostatin, CD43)	20.60		
134182	D52059	Hs.7972	KIAA0871 protein	12.20		
133985	L34657	Hs.78146	platelet/endothelial cell adhesion molec			
133978	AF035718	Hs.78061	transcription factor 21			
133835	AI677897	Hs.76640	RGC32 protein			
133651	AI301740	Hs.173381	dihydropyrimidinase-like 2			
133633	D21252	Hs.75337	nucleolar and coiled-body phosphoprotein	15.20		
133565	AW955776	Hs.313500	ESTs, Moderately similar to ALL7_HUMAN A			
133548	AW646384	Hs.178112	DNA segment, single copy probe LNS-CAVL			1.77
133488	AA335295	Hs.74120	adipose specific 2			
133478	X83703	Hs.31432	cardiac ankyrin repeat protein			2.08
133337	AF085983	Hs.293676	ESTs		9.60	
133200	AB037715	Hs.183639	hypothetical protein FLJ10210			1.77
133153	AF070592	Hs.66170	HSKM-B protein	30.60		
133130	AI126606	Hs.6557	zinc finger protein 181	22.60		
133120	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein			
132928	AW168082	Hs.169449	protein kinase C, alpha	13.80		
132836	AB023177	Hs.29900	KIAA0960 protein			
132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,	41.60		
132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypotheti	40.40		
132548	X12830	Hs.193400	interleukin 6 receptor		7.20	
132476	AL119844	Hs.49476	Homo sapiens clone TUA8 Crl-du-chat regi		4.75	
132439	AK001942	Hs.4863	hypothetical protein DKFZp566A1524			1.88
132240	AB018324	Hs.42676	KIAA0781 protein	21.20		
132210	NM_007203	Hs.42322	A kinase (PRKA) anchor protein 2			1.99
132199	AL041299	Hs.165084	ESTs	15.20		
131751	T96555	Hs.31562	ESTs			1.76
131745	AI828559	Hs.31447	ESTs, Moderately similar to A46010 X-U	27.80		
131694	NM_000246	Hs.3076	MHC class II transactivator		4.00	
131688	NM_012296	Hs.30687	GRB2-associated binding protein 2			
131676	AI126821	Hs.30514	ESTs		6.20	
131629	Z45794	Hs.238809	ESTs	21.40		
131589	C18825	Hs.29191	epithelial membrane protein 2			
131536	AA019201	Hs.269210	ESTs		9.40	
131517	AB037789	Hs.263395	sema domain, transmembrane domain (TM),		3.59	
131355	R52804	Hs.25956	DKFZP564D206 protein		4.48	
131253	R71802	Hs.24853	ESTs	15.00		
131207	AF104266	Hs.24212	latrophilin			1.75
131156	AI472209	Hs.323117	ESTs			1.84
131066	AW169287	Hs.22588	ESTs		3.54	
131061	N64328	Hs.268744	KIAA1795 protein			
131053	AA348541	Hs.296261	guanine nucleotide binding protein (G pr			1.93
130895	AA641767	Hs.21015	hypothetical protein DKFZp564L0864 simil	16.60		
130762	D84371	Hs.1898	paraoxonase 1	12.00		

	130657	AW337575	Hs.201591	ESTs		
	130655	AI831962	Hs.17409	cysteine-rich protein 1 (intestinal)		
	130589	AL110226	Hs.16441	DKFZP434H204 protein		2.08
5	130562	D50402	Hs.182611	solute carrier family 11 (proton-coupled)		1.91
	130555	R69743	Hs.116774	integrin, alpha 1	9.60	
	130365	W56119	Hs.155103	eukaryotic translation initiation factor	11.60	
	130273	AW972422	Hs.153863	MAD (mothers against decapentaplegic, Drosophila)	6.60	
	130259	NM_000328	Hs.153614	refinitis pigmentosa GTPase regulator		1.91
10	130090	H97878	Hs.132390	zinc finger protein 36 (KOX 18)	21.20	
	129958	R27496	Hs.1378	annexin A3	5.05	
	129898	AI672731	Hs.13256	ESTs		
	129875	AA181018	Hs.13055	hypothetical protein FLJ13920	18.60	
	129699	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligase		
15	129626	F13272	Hs.111334	ferritin, light polypeptide		
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	22.63	
	129593	AI338247	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f)		
	129565	X77777	Hs.198726	vasoactive intestinal peptide receptor 1		2.53
	129527	AA769221	Hs.270847	delta-tubulin	39.20	
20	129402	W72052	Hs.111112	ESTs		2.11
	129385	AA172106	Hs.110950	Rag C protein	15.20	
	129315	NM_014563	Hs.174038	spondyloepiphyseal dysplasia, late	12.40	
	129312	T97579	Hs.110334	ESTs, Weakly similar to I78885 serine/threonine	20.83	
	129240	AA361258	Hs.237868	interleukin 7 receptor		1.95
	129210	AL039940	Hs.202949	KIAA1102 protein		
25	129122	AW958473	Hs.301957	nucleoside diphosphate linked moiety X motif 2	4.20	
	129057	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)		
	128946	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3-monooxygenase)	5.20	
	128798	AF015525	Hs.302043	chemokine (C-C motif) receptor-like 2		
30	128789	AW368576	Hs.139851	caveolin 2		2.24
	128778	AA504776	Hs.186709	ESTs, Weakly similar to I38022 hypothetical protein	12.20	
	128766	AW160432	Hs.296460	craniofacial development protein 1	26.40	
	128631	R44238	Hs.155546	KIAA1080 protein; Golgi-associated, gamma		1.78
	128624	BE154765	Hs.102647	ESTs, Weakly similar to TRHY_HUMAN TRICH		2.51
35	128609	NM_003616	Hs.102456	survival of motor neuron protein interactor	16.00	
	128603	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHITE)	12.80	
	128598	AA305407	Hs.102308	potassium inwardly-rectifying channel, subfamily A member 1	4.00	
	128458	H55864	Hs.56340	ESTs		
40	128051	AF150882	Hs.186877	sodium channel, voltage-gated, type XII, alpha	17.20	
	127958	AA830201	Hs.124347	ESTs	21.30	
	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L		
	127944	AI557081	Hs.262476	S-adenosylmethionine decarboxylase 1	10.60	
	127825	AA805151	Hs.3628	mitogen-activated protein kinase kinase	13.40	
	127896	AI669586	Hs.222194	ESTs		7.00
45	127859	AA761802	Hs.291559	ESTs	14.00	
	127817	AA836641	Hs.163085	ESTs	14.00	
	127742	AW293496	Hs.180138	ESTs	11.00	
	127628	AI240102	Hs.322430	NDRG family, member 4	11.10	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture)		
50	127582	AA908954	Hs.130844	ESTs	19.60	
	127543	AK000787	Hs.157392	Homo sapiens cDNA FLJ20780 fis, clone CO	15.40	
	127535	AA568424	Hs.164450	ESTs	17.50	
	127404	AI379920	Hs.270224	ESTs	14.80	
	127396	L31968	Hs.187991	DKFZP564A122 protein	15.40	
55	127374	AA442797	Hs.312110	ESTs, Weakly similar to I38022 hypothetical protein	14.60	
	127346	AA203616	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, member 1	21.00	
	127340	BE047653	Hs.119183	ESTs, Weakly similar to ZN91_HUMAN ZINC	15.80	
	127307	AW952712	Hs.126712	ESTs, Weakly similar to AF191020 1 E2IG5		
	127242	AW390395	Hs.181301	cathepsin S	22.60	
60	127167	AA625690	Hs.190272	ESTs	21.40	
	127046	AA321948	Hs.293968	ESTs	41.20	
	126928	AA480902	Hs.137401	ESTs	11.00	
	126900	AF137386	Hs.12701	plasmalogen		1.78
	126852	AA399961		gb:zu68c01.r1 Soares_testis_NHT Homo sap	5.60	
65	126816	AA248234		gb:csq2228.seq.F Human fetal heart, Lamb	12.20	
	126812	AB037860	Hs.173933	nuclear factor I/A	17.19	
	126666	AA648886	Hs.151999	ESTs	13.57	
	126645	AA316181	Hs.61635	six transmembrane epithelial antigen of	15.40	
	126592	AI611153	Hs.6093	Homo sapiens cDNA: FLJ22783 fis, clone K		4.67
70	126556	AF255303	Hs.112227	membrane-associated nucleic acid binding protein	18.00	
	126433	AA325606		gb:EST28707 Cerebellum II Homo sapiens c	16.77	
	126299	AW979155	Hs.298275	amino acid transporter 2	14.60	
	126218	AL049801	Hs.13649	Novel human gene mapping to chromosome 13		3.50
	126182	AA721331	Hs.293771	ESTs	13.40	
75	126177	AW752782	Hs.129750	hypothetical protein FLJ10546	18.20	
	126142	H86261	Hs.40568	ESTs	14.00	
	126077	M78772	Hs.210836	ESTs	16.59	
	125994	AI990529	Hs.270799	ESTs	17.40	
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	13.00	
80	125847	AW161885	Hs.249034	ESTs	49.57	
	125831	HO4043		gb:yl45c03.r1 Soares placenta Nb2HP Homo		
	125731	R61771	Hs.26912	ESTs	13.20	
	125676	BE612918	Hs.151973	hypothetical protein FLJ23511	11.20	
	125561	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S		
85	125552	HO9701	Hs.278366	ESTs, Weakly similar to I38022 hypothetical protein	12.60	
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	33.40	

	125422	AA903229	Hs.153717	ESTs		1.80
	125331	AI422996	Hs.161378	ESTs	38.00	
	125309	T12411	Hs.183745	hypothetical protein FLJ13456	18.20	
5	125167	AL137540	Hs.102541	netrin 4		1.95
	125139	AW194933	Hs.9788	hypothetical protein MGC10924 similar to		1.84
	125042	T78906	Hs.269432	ESTs, Moderately similar to ALU1_HUMAN	21.80	
	124711	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	10.60	
	124631	NM_014053	Hs.270594	FLVCR protein	23.20	
10	124578	N68321	Hs.231500	EST	21.43	
	124574	AL036596	Hs.42322	A kinase (PRKA) anchor protein 2		1.77
	124472	N52517	Hs.102670	EST	37.20	
	124438	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A		
	124357	N22401		gb:yyw37g07.s1 Morton Fetal Cochlea Homo	14.64	
15	124306	AW973078	Hs.293039	ESTs		4.00
	124214	H58608	Hs.151323	ESTs		
	124097	AW298235	Hs.101689	ESTs	27.20	
	123978	T89832	Hs.170278	ESTs		2.03
	123972	T46848	Hs.70337	immunoglobulin superfamily, member 4	6.00	
20	123961	AL050184	Hs.21610	DKFZP434B203 protein		1.79
	123936	NM_004673	Hs.241619	angiotensin-like 1	15.80	
	123802	AA620448		gb:ae58c09.s1 Stratagene lung carcinoma	4.23	
	123734	AA609861	Hs.312447	ESTs	4.20	
25	123619	AA602964		gb:mo97c02.s1 NCL_CGAP_Pr2 Homo sapiens	33.60	
	123596	AA421130	Hs.112640	EST	10.93	
	123476	AA384664	Hs.108829	ESTs		2.18
	123340	AA504264	Hs.182937	peptidylprolyl isomerase A (cyclophilin	11.20	
	123190	AA489212	Hs.105228	EST	14.20	
	123136	AW451999	Hs.194024	ESTs		7.00
30	123073	AA485061	Hs.105652	ESTs	31.20	
	123055	AA482005	Hs.105102	ESTs, Weakly similar to reverse transcri		4.80
	122699	AA456130	Hs.301721	KAA1255 protein		5.00
	122679	AA811286	Hs.192837	ESTs, Weakly similar to ALU5_HUMAN ALU S	14.40	
	122633	NM_001546	Hs.34853	inhibitor of DNA binding 4, dominant neg		
35	122553	AA451884	Hs.190121	ESTs	40.00	
	122544	AW973253	Hs.292689	ESTs	15.40	
	122485	AA524547	Hs.160318	FXID domain-containing ion transport reg		1.81
	122211	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 H	12.10	
	122127	AW207175	Hs.106771	ESTs		1.95
40	122011	AA431082		gb:zw78a10.s1 Soares_testis_NHT Homo sap		1.89
	121992	AI860775	Hs.98506	ESTs	3.60	
	121989	W56487	Hs.193784	Homo sapiens mRNA; cDNA DKFZp586K1922 (f		2.01
	121835	AB033030	Hs.300670	KAA1204 protein		1.85
	121726	AF241254	Hs.178098	angiotensin I converting enzyme (peptidyl	12.43	
45	121690	AV660305	Hs.110286	ESTs		1.82
	121643	AA640987	Hs.193767	ESTs		
	121633	AA417011	Hs.98175	EST	14.00	
	121622	AA416931	Hs.126065	ESTs		16.40
	121497	AA412031	Hs.97901	EST	11.20	
50	121351	AW206227	Hs.287727	hypothetical protein FLJ23132	12.20	
	121314	W07343	Hs.182538	phospholipid scramblase 4		1.83
	121242	AA400857	Hs.97609	ESTs	22.40	
	121059	AA393283		gb:zt74e03.r1 Soares_testis_NHT Homo sap	14.80	
	120934	AA226198		gb:nc26a07.s1 NCL_CGAP_Pr1 Homo sapiens	21.20	
55	120755	AA312934	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone		1.79
	120637	AA811804		gb:nb39a05.s1 NCL_CGAP_GCB1 Homo sapiens	20.00	
	120484	AA253170	Hs.95473	EST	40.20	
	120336	N85785	Hs.181165	eukaryotic translation elongation factor		6.60
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypotheti	16.80	
60	120132	W57554	Hs.125019	ESTs		4.73
	120041	AA830882	Hs.59368	ESTs		1.75
	119996	W88996	Hs.59134	EST		7.20
	119970	AA767718	Hs.93581	hypothetical protein FLJ10512	11.20	
	119861	W78816	Hs.49943	ESTs, Weakly similar to S65657 alpha-1C-		3.78
65	119824	W74536	Hs.184	advanced glycosylation end product-speci		
	119740	AW021407	Hs.21068	hypothetical protein	20.20	
	119271	AI061118	Hs.65328	Fanconi anemia, complementation group F	15.20	
	119221	C14322	Hs.250700	tryptase beta 1		
	119126	R45175	Hs.117183	ESTs	12.60	
70	119073	BE245360	Hs.279477	ESTs		
	118928	AA312799	Hs.283689	activator of CREM in testis		10.00
	118901	AW292577	Hs.94445	ESTs		3.96
	118661	AL137554	Hs.49927	protein kinase NYD-SP15		9.60
	118607	AI377444	Hs.54245	ESTs, Weakly similar to S65824 reverse t	10.40	
75	118449	AI813865	Hs.164478	hypothetical protein FLJ21939 similar to		1.90
	118416	N66028	Hs.49105	FKBP-associated protein	16.20	
	118379	N64491	Hs.48990	ESTs		4.00
	118329	N63520		gb:yyg6201.s1 Soares_multiple_sclerosis_		6.60
	118320	N63451	Hs.141600	ESTs, Weakly similar to alternatively s		3.80
80	118253	AA497044	Hs.20887	hypothetical protein FLJ10392	17.60	
	118124	N66968	Hs.46707	chromosome 21 open reading frame 37	14.00	
	118055	AB037746	Hs.42768	hypothetical protein DKFZp761O0113		1.86
	118032	N52802	Hs.47544	EST		5.00
	117840	T26379	Hs.48802	Homo sapiens clone 23632 mRNA sequence	4.00	
85	117404	N39725	Hs.15220	zinc finger protein 106		1.90
	117314	N32498	Hs.42829	ESTs	14.20	

	117209	W03011	Hs.306881	MSTP043 protein			
	117023	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	2.31		
	116814	H50834		gb:yp86a10.s1 Soares fetal liver spleen	20.20		
5	116784	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	3.51		
	116766	AI608657	Hs.95097	ESTs	16.20		
	116712	AW901618	Hs.61935	Homo sapiens mRNA; cDNA DKFZp7611071 (fr	6.80		
	116707	H10344	Hs.49050	ESTs, Weakly similar to A Chain A, Human	18.60		
	116351	AL133623	Hs.82501	similar to mouse Xn1 / Dhml2 protein	19.40		
10	116279	AW971248	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S			
	116166	AL039940	Hs.202949	KIAA1102 protein	2.13		
	116152	AL040521	Hs.15220	zinc finger protein 106	1.75		
	116117	BE613410	Hs.31575	SEC63, endoplasmic reticulum translocon	13.20		
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	30.11		
15	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	2.36		
	115955	AF263613	Hs.44198	intracellular membrane-associated calciu	18.20		
	115844	AI373062	Hs.332938	hypothetical protein MGC5370	18.57		
	115683	AF255910	Hs.54650	junctional adhesion molecule 2	23.00		
	115673	AA406341	Hs.269908	Homo sapiens cDNA FLJ11991 fis, clone HE	11.82		
20	115672	AI889110	Hs.73251	ESTs	10.60		
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1	1.76		
	115313	AA808001	Hs.184411	albumin	25.20		
	115279	AW964897	Hs.290825	ESTs	8.00		
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	1.80		
25	115110	AK001671	Hs.11387	KIAA1453 protein	14.20		
	114999	BE246481	Hs.87856	ESTs	19.20		
	114930	AA237022	Hs.188717	ESTs	5.60		
	114922	AA235672	Hs.87491	ESTs	3.60		
	114837	BE244930	Hs.166895	ESTs	43.70		
30	114769	AA149060	Hs.296100	ESTs	11.00		
	114761	AA143781	Hs.126280	hypothetical protein FLJ23393	14.00		
	114736	AB10347	Hs.103812	ESTs, Moderately similar to ALU1_HUMAN A	4.20		
	114596	AA310162	Hs.169248	cytochrome c	10.71		
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	20.40		
35	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.40		
	114452	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE	17.20		
	114359	NM_016929	Hs.283021	chloride intracellular channel 5	2.09		
	114357	R41677	Hs.6107	Homo sapiens cDNA FLJ14839 fis, clone OV	12.40		
	114251	H15261	Hs.21948	ESTs	2.00		
40	114138	AW384793	Hs.15740	Homo sapiens mRNA; cDNA DKFZp434E033 (fr	11.40		
	114124	W57554	Hs.125019	ESTs	6.04		
	113946	AW083883	Hs.37896	Homo sapiens cDNA FLJ13510 fis, clone PL	1.82		
	113695	T96965	Hs.17948	ESTs, Weakly similar to ALU8_HUMAN IIII	2.15		
	113606	NM_013343	Hs.278951	NAG-7 protein			
45	113590	R49642	Hs.142447	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.60		
	113560	T91015	Hs.268626	ESTs	32.00		
	113552	AI654223	Hs.16026	hypothetical protein FLJ23191			
	113540	AW152618	Hs.16757	ESTs			
	113502	T89130		gb:ye12d01.s1 Stratagene lung (937210) H	8.35		
50	113288	AI076838	Hs.12967	ESTs	12.40		
	113252	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	4.27		
	113238	R45467	Hs.189813	ESTs			
	113203	AA743563	Hs.10305	ESTs	21.20		
	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	1.92		
55	113089	T40707	Hs.270862	ESTs	14.33		
	113076	AF033199	Hs.81198	zinc finger protein 204	6.00		
	113009	T23699	Hs.7246	ESTs	9.40		
	112937	AI694320	Hs.6295	ESTs, Weakly similar to T17248 hypotheti	12.20		
	112891	T03927	Hs.293147	ESTs, Moderately similar to A46010 X-li	10.57		
60	112794	R97018		gb:cyq74b08.s1 Soares fetal liver spleen	26.60		
	112691	R88708	Hs.220647	ESTs	15.33		
	112602	AW004045	Hs.203365	ESTs	15.60		
	112366	AF035318	Hs.12533	Homo sapiens clone 23705 mRNA sequence	15.40		
	112210	R49645	Hs.7004	ESTs	14.00		
65	112064	AL049390	Hs.22689	Homo sapiens mRNA; cDNA DKFZp586O1318 (f	13.00		
	111998	R42379	Hs.138283	ESTs	11.00		
	111987	NM_015310	Hs.6763	KIAA0942 protein	22.40		
	111803	AA593731	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.77		
	111737	H04607	Hs.9218	ESTs	1.86		
70	111605	T91081	Hs.194178	ESTs, Moderately similar to PC4259 ferri	23.00		
	111510	R07856	Hs.16355	ESTs	11.02		
	111341	AL157484	Hs.22483	Homo sapiens mRNA; cDNA DKFZp762M127 (fr	1.88		
	111280	AA373527	Hs.19385	CGI-58 protein	18.40		
	111247	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f			
75	111232	AI247763	Hs.16928	ESTs	27.60		
	110942	R63503	Hs.28419	ESTs	14.80		
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	24.71		
	110837	H03109	Hs.108920	HT018 protein	2.18		
	110824	AI767183	Hs.26942	ESTs	12.20		
80	110776	AB032417	Hs.19545	frizzled (Drosophila) homolog 4	1.75		
	110576	H60869	Hs.37889	ESTs	13.00		
	110369	AK000768	Hs.107872	hypothetical protein FLJ20761	5.60		
	110099	R44557	Hs.23748	ESTs	2.31		
	109984	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL			
85	109958	AA001266	Hs.133521	ESTs	11.25		
	109893	AA884208	Hs.30484	ESTs	2.68		



5	109842	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	23.83		
	109837	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypothe		3.91	
	109796	A1800515	Hs.12024	ESTs		17.20	
	109688	R41900	Hs.22245	ESTs		9.60	
	109648	H17800	Hs.7154	ESTs	22.80		
10	109613	H47315	Hs.27519	ESTs			
	109550	AW021488	Hs.26981	ESTs			
	109523	AW193342	Hs.24144	ESTs		1.89	
	109472	AK001989	Hs.91165	hypothetical protein		6.00	
	109355	AA524525	Hs.48297	DKFZP586C1620 protein	15.00		
15	109260	AW978515	Hs.131915	KIAA0863 protein	25.60		
	108781	AA128654		gbzn98g07.s1 Stratagene fetal retina 93	14.20		
	108663	BE219231	Hs.292653	ESTs, Weakly similar to T26845 hypothe	11.00		
	108573	AA086005		gbzl84c04.s1 Stratagene colon (937204)	26.00		
	108480	AL133092	Hs.68055	hypothetical protein DKFZp434i0428			
20	108382	NM_006770	Hs.67726	macrophage receptor with collagenous str		1.83	
	108174	AA055632	Hs.303070	ESTs	15.20		
	108138	ALD49990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr		3.60	
	108087	AA045708	Hs.40545	ESTs	15.44		
	108048	AI797341	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT		11.40	
25	108041	AW204712	Hs.61957	ESTs			
	107997	ALD49176	Hs.82223	chordin-like		4.76	
	107994	AA036811	Hs.48469	LIM domains containing 1			
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	14.20		
	107681	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN A	51.80		
30	107666	AA010511	Hs.60418	EST	29.20		
	107332	T87750	Hs.183297	DKFZP566F2124 protein	10.73		
	107292	BE166479	Hs.4789	Homo sapiens serologically defined breas	32.00		
	107230	AI034467	Hs.34650	ESTs	17.40		
	107168	W57578	Hs.237955	RAB7, member RAS oncogene family	10.43		
35	107160	AA314490	Hs.27669	KIAA1563 protein	11.40		
	107054	AI076459	Hs.15978	KIAA1272 protein			
	107029	AF264750	Hs.288971	myeloid/lymphoid or mixed-lineage leukem	21.40		
	106999	H93281	Hs.10710	hypothetical protein FLJ20417	35.80		
	106954	AF128847	Hs.204038	indolethylamine N-methyltransferase		1.76	
40	106870	AI983730	Hs.26530	serum deprivation response (phosphatidyl			
	106865	AW192535	Hs.19479	ESTs	13.40		
	106844	AA485055	Hs.158213	sperm associated antigen 6		7.13	
	106820	NM_016831	Hs.12592	period (Drosophila) homolog 3		7.00	
	106818	AK002135	Hs.3542	hypothetical protein FLJ11273	13.00		
45	106797	AI768801	Hs.169943	Homo sapiens cDNA FLJ13569 fis, clone PL		2.05	
	106773	AA478109	Hs.188833	ESTs			
	106747	NM_007118	Hs.171957	triple functional domain (PTPRF interact	12.60		
	106743	BE613328	Hs.21938	hypothetical protein FLJ12492	10.60		
	106687	AW360847	Hs.16578	ESTs			
50	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr		2.40	
	106567	AW450408	Hs.86412	chromosome 9 open reading frame 5		1.78	
	106562	ALD31846	Hs.152151	plakophilin 4		1.76	
	106536	AA328648	Hs.23804	ESTs, Weakly similar to PN0099 son3 prot		2.19	
	106533	AL134708	Hs.145998	ESTs	23.20		
55	106507	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	15.20		
	106490	AA404255	Hs.115537	putative dipeptidase			
	106474	BE383658	Hs.42484	hypothetical protein FLJ10618	10.44		
	106211	AA428240	Hs.126083	ESTs		29.80	
	105986	AB037722	Hs.8707	KIAA1301 protein		3.70	
60	105894	AI904740	Hs.25691	receptor (calcitonin) activity modifying		1.94	
	105847	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-		1.75	
	105803	AW747996	Hs.160999	ESTs, Moderately similar to A56194 throm		2.47	
	105731	AA834664	Hs.29131	nuclear receptor coactivator 2	10.71		
	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds			
65	105688	AI299139	Hs.17517	ESTs	23.40		
	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	37.20		
	105101	H53202	Hs.38163	ESTs		8.30	
	104989	R65998	Hs.285243	hypothetical protein FLJ22029		8.09	
	104986	AW088826	Hs.117176	poly(A)-binding protein, nuclear 1		1.92	
70	104969	AI670947	Hs.78406	phosphatidylinositol-4-phosphate 5-kinas		5.40	
	104903	AI436323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,		7.60	
	104896	AW015318	Hs.23165	ESTs	13.80		
	104865	T79340	Hs.22575	Homo sapiens cDNA: FLJ21042 fis, clone C			
	104825	AA035613	Hs.141883	ESTs		1.87	
75	104781	AA099904	Hs.21610	DKFZP434B203 protein		1.93	
	104776	AA026349		gb:z99f01.s1 Soares_pregnant_uterus_NbH	10.20		
	104691	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor		5.69	
	104667	AI239923	Hs.30098	ESTs		3.82	
	104404	H58762		gb:EST00057 HE6W Homo sapiens cDNA clone	4.20		
80	104392	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	27.20		
	104212	AB002298	Hs.173035	KIAA0300 protein		1.91	
	104074	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	11.20		
	103749	AL135301	Hs.8768	hypothetical protein FLJ10849	10.66		
	103645	AW246253	Hs.7043	succinate-CoA ligase, GDP-forming, alpha	12.00		
85	103554	AI878826	Hs.323469	caveolin 1, caveolae protein, 22kD		1.80	
	103541	AI815801	Hs.79197	CD83 antigen (activated B lymphocytes, I			
	103496	Y09267	Hs.132821	flavin containing monooxygenase 2			
	103428	BE383507	Hs.78921	A kinase (PRKA) anchor protein 1	11.20		
	103353	X89399	Hs.119274	RAS p21 protein activator (GTPase activa	19.80		

	103295	X81479	Hs.2375	egf-like module containing, mucin-like,	3.60	
	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula		
	103100	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)		1.76
5	103025	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t		2.15
	102698	M18667	Hs.1867	progastrin (pepsinogen C)		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein	11.00	
	102580	U60808	Hs.152981	CDP-diacylglycerol synthase (phosphatida	25.40	
	102417	AA034127	Hs.153487	signal transducing adaptor molecule (SH3	14.00	
10	102363	NM_003734	Hs.198241	amine oxidase, copper containing 3 (vasc		
	102302	AA306342	Hs.69171	protein kinase C-like 2	10.86	
	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11		
	102188	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	7.40	
	102151	T27013	Hs.3132	steroidogenic acute regulatory protein	16.40	
15	101957	L28824	Hs.74101	spleen tyrosine kinase	15.40	
	101842	M93221	Hs.75182	mannose receptor, C type 1		
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant		
	101764	A198550	Hs.81256	S100 calcium-binding protein A4 (calcium		1.78
	101716	AF050558	Hs.2563	tachykinin, precursor 1 (substance K, su	18.80	
20	101678	M62505	Hs.2161	complement component 5 receptor 1 (C5a)		2.22
	101447	M21305		gb:Human alpha satellite and satellite 3	504.80	
	101383	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	31.00	
	101348	A1738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N		1.75
	101345	NM_005795	Hs.152175	calcitonin receptor-like		
25	101336	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h		2.24
	101330	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do		
	101277	BE297626	Hs.296049	microfibrillar-associated protein 4		
	101262	L35854		gb:Human dystrophin (dp140) mRNA, 5' end	19.00	
	101168	NM_005308	Hs.211559	G protein-coupled receptor kinase 5		2.01
30	101102	NM_003243	Hs.79059	transforming growth factor, beta recepto		
	101088	X70697	Hs.553	solute carrier family 6 (neurotransmitte	7.52	
	101066	AW970254	Hs.889	Charot-Leyden crystal protein	19.38	
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte		1.91
	100893	BE245294	Hs.180789	S164 protein	15.40	
35	100770	W25797.comp	Hs.177486	amyloid beta (A4) precursor protein (pro	11.20	
	100716	X89887	Hs.172350	HIR (histone cell cycle regulation defec	14.80	
	100555	M69181		gb:Human nonmuscle myosin heavy chain-B	33.00	
	100425	NM_014747	Hs.78748	KIAA0237 gene product	16.20	
	100408	D66540	Hs.56045	src homology three (SH3) and cysteine ri	4.00	
40	100382	D83407	Hs.156007	Down syndrome critical region gene 1-lik	4.24	
	100351	D64158			6.20	
	100299	D49493	Hs.2171	growth differentiation factor 10	21.20	
	100134	AA305746	Hs.49	macrophage scavenger receptor 1		
	100108	U09577	Hs.76873	hyaluronoglucosaminidase 2		1.79
45	100095	Z97171	Hs.78454	myocilin, trabecular meshwork inducible	5.40	
	100066				11.29	

TABLE 3B shows the accession numbers for those primekeys lacking unigenelD's for Table 3A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325606 AA099517 N89423
125831	1522905_1	H04043 D60988 D60337
126816	122973_1	AA248234 AA090985
65	126852	136135_1 AA399961 AA128347
	121059	273450_1 AA393283 AA398628
	120537	200885_1 AAB11804 AAB09404 AA286907 AW977624
	122011	7617_-2 AA431082
70	120934	177521_1 AA226198 AA226513 AA383773
	123802	genbank_AA620448 AA620448
	116814	genbank_H50834 H50834
	118329	genbank_N63520 N63520
	104404	H58762_at H58762
	104776	genbank_AA026349 AA026349
75	113502	genbank_T89130 T89130
	101262	entrez_L35854 L35854
	108573	genbank_AA086005 AA086005
	101447	entrez_M21305 M21305
	124357	genbank_N22401 N22401
80	108781	genbank_AA128654 AA128654
	112794	genbank_R97018 R97018
	100351	entrez_D64158 D64158
85	100555	tigr_HT2245 M69181 M81105 U51039

Table 4A shows 202 genes up-regulated in samples from patients treated with chemotherapy or radiotherapy. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5 Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: average of AI for samples from patients treated with chemotherapy or radiotherapy divided by the average of AI for normal lung samples.

	Pkey	ExAccn	UnigeneID	Unigene Title	R1
	100113	NM_001269	Hs.84746	chromosome condensation 1	27.20
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	20.60
15	100210	D26361	Hs.3104	KIAA0042 gene product	20.40
	100225	D28539	Hs.167185	glutamate receptor, metabotropic 5	20.60
	100269	NM_001949	Hs.1189	E2F transcription factor 3	29.40
	100438	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	23.50
	100877	X80821	Hs.27973	KIAA0874 protein	35.56
20	100893	BE245294	Hs.180789	S164 protein	43.40
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	21.80
	101447	M21305		gb:Human alpha satellite and satellite 3	193.60
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	38.40
	101724	L11650	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	198.80
25	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	78.60
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	162.20
	101879	AA176374	Hs.243886	nuclear autoantigenic sperm protein (his	50.00
	101915	AF207881	Hs.155185	cytosolic ovarian carcinoma antigen 1	26.00
	101973	U41514	Hs.80120	UDP-N-acetyl-alpha-D-galactosamine:polyp	37.20
30	102025	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,	
	102031	U04898	Hs.2156	RAR-related orphan receptor A	32.00
	102052	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	51.20
	102391	AA296874	Hs.77494	deoxyguanosine kinase	13.90
	102420	U44060	Hs.14427	Homo sapiens cDNA: FLJ21800 fis, clone H	28.80
35	102610	U65011	Hs.30743	preferentially expressed antigen in mela	110.60
	102829	NM_006183	Hs.80962	neurotensin	116.80
	103000	NM_001975	Hs.146580	enolase 2, (gamma, neuronal)	2.30
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	181.40
	103507	AJ000512	Hs.296323	serum/glucocorticoid regulated kinase	49.20
40	103587	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein	88.60
	104660	BE298665	Hs.14846	Homo sapiens mRNA: cDNA DKFZp564D016 (fr	42.60
	104896	AW015318	Hs.23165	ESTs	29.40
	105038	AW503733	Hs.9414	KIAA1488 protein	21.50
	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	32.80
45	105510	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	20.20
	105567	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specifi	28.40
	106073	AL157441	Hs.17834	downstream neighbor of SON	25.40
	106205	AW965058	Hs.111583	ESTs, Weakly similar to I38022 hypotheri	32.00
	106516	AL137311	Hs.234074	Homo sapiens mRNA: cDNA DKFZp761G02121 (	40.60
50	106533	AL134708	Hs.145998	ESTs	59.80
	106575	AW970602	Hs.105421	ESTs	43.40
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	50.80
	106851	AJ458623		gb:cd04g09.x1 NCI_CGAP_Lu24 Homo sapiens	53.40
	106995	AB023139	Hs.37892	KIAA0922 protein	20.88
55	107332	T87750	Hs.183297	DKFZP566F2124 protein	23.60
	107532	AA443473	Hs.173694	Homo sapiens mRNA: cDNA DKFZp762G207 (fr	57.20
	107922	BE153855	Hs.61460	Ig superfamily receptor LNIR	49.00
	108609	BE409857	Hs.69499	hypothetical protein	19.67
	108780	AL076442	Hs.117938	collagen, type XVII, alpha 1	48.17
60	109166	AA219691	Hs.73625	RAB6 interacting, kinesin-like (raklins	59.20
	109260	AW978515	Hs.131915	KIAA0863 protein	28.60
	109280	AK001355	Hs.279610	hypothetical protein FLJ10493	22.80
	109292	AW975746	Hs.188662	KIAA1702 protein	
65	109384	AA219172	Hs.86849	ESTs	21.00
	109415	U80736	Hs.110826	trinucleotide repeat containing 9	31.60
	109445	AA232103	Hs.189915	ESTs	24.20
	109502	AW967069	Hs.211556	hypothetical protein MGC5487	21.40
	109633	AW003785	Hs.170267	ESTs	20.40
	109786	AI989482	Hs.146286	kinesin family member 13A	19.60
70	109958	AA001266	Hs.133521	ESTs	24.00
	110920	N47224	Hs.20521	HMT1 (hnRNP methyltransferase, S. cerevi	28.40
	110924	AW058463	Hs.12940	zinc-fingers and homeoboxes 1	35.00
	111084	H44186	Hs.15456	PDZ domain containing 1	61.20
	111132	AB037807	Hs.83293	hypothetical protein	24.60
75	111229	AW389845	Hs.110855	ESTs	27.20
	111337	AA837396	Hs.263925	US1-interacting protein NUDE1, rat homo	48.00
	111887	NM_015310	Hs.6763	KIAA0942 protein	37.80
	112048	AA383343	Hs.22116	CDC14 (cell division cycle 14, S. cerevi	25.80
	112268	W39509	Hs.22003	solute carrier family 6 (neurotransmitte	63.80
80	112685	R87650	Hs.33439	ESTs, Weakly similar to ALU1_HUMAN ALU	26.40
	112671	AL110216	Hs.12285	ESTs, Weakly similar to I55214 salivary	47.64
	112697	AW206453	Hs.3782	ESTs	22.00
	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	65.00
	112992	AL157425	Hs.133315	Homo sapiens mRNA: cDNA DKFZp761J1324 (f	42.00
85	113073	N39342	Hs.103042	microtubule-associated protein 1B	55.40

	113494	T91451	Hs.86538	ESTs	22.80
	113560	T91015	Hs.268626	ESTs	22.80
	113849	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	51.80
5	113950	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	28.20
	114339	AA782845	Hs.22790	ESTs	20.20
	114365	H42169	Hs.18653	hypothetical protein FLJ14627	21.00
	114455	H37908	Hs.271616	ESTs, Weakly similar to ALU8_HUMAN ALU S	25.80
	114518	AW163267	Hs.106469	suppressor of var1 (S.cerevisiae) 3-like	23.60
10	114824	AA950951	Hs.305953	zinc finger protein 83 (HPF1)	27.20
	114837	BE244930	Hs.166895	ESTs	30.20
	114974	AW956931	Hs.179662	nucleosome assembly protein 1-like 1	20.80
	115075	AA814043	Hs.88045	ESTs	30.60
	115084	BE383668	Hs.42484	hypothetical protein FLJ10618	28.86
15	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	38.00
	115313	AA808001	Hs.184411	albumin	22.60
	115697	D31382	Hs.63325	transmembrane protease, serine 4	173.60
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	27.77
	116090	AI591147	Hs.61232	ESTs	20.80
20	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	164.20
	116399	AA889120	Hs.110637	homeo box A10	38.00
	117099	H93699		gb:yy16a11.s1 Soares fetal liver spleen	21.60
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	49.40
	118091	AW005054	Hs.47883	ESTs, Weakly similar to KCC1_HUMAN CALCI	22.40
25	118138	AA374756	Hs.93560	Homo sapiens mRNA for KIAA1771 protein,	22.00
	118720	N73515		gb:ze49d07.s1 Soares fetal liver spleen	20.00
	118873	AI824009	Hs.44577	ESTs	19.40
	119126	R45175	Hs.117183	ESTs	111.20
	119717	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	33.00
30	119940	AL050097	Hs.272531	DKFZP586B0319 protein	31.00
	120266	AI807264	Hs.205442	ESTs, Weakly similar to T34036 hypothe	20.20
	120515	AA258356		gb:zr59c10.s1 Soares_NhHMPu_S1 Homo sapi	25.00
	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	95.40
	120983	AA398209	Hs.97587	EST	105.20
35	121054	AW976570	Hs.97387	ESTs	38.80
	121369	AW450737	Hs.128791	CGI-09 protein	41.60
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	30.80
	122612	AA974832	Hs.128708	ESTs	19.60
	123130	AA487200		gb:ab19f02.s1 Stralagene lung (937210) H	33.20
40	123440	AI733692	Hs.112488	ESTs	23.17
	123596	AA421130	Hs.112640	EST	23.00
	123619	AA602964		gb:nc97c02.s1 NCL_CGAP_Pr2 Homo sapiens	28.80
	124006	AI147155	Hs.270016	ESTs	77.60
	124169	BE079334	Hs.271630	ESTs	22.20
45	124281	AI333756	Hs.111801	arsenate resistance protein ARS2	42.20
	124472	N52517	Hs.102670	EST	32.60
	124617	AW628168	Hs.152684	ESTs	21.80
	124631	NM_014053	Hs.270594	FLVCR protein	30.40
	124839	R55784	Hs.140942	ESTs	21.20
50	125186	AA610620	Hs.181244	major histocompatibility complex, class	42.80
	125321	T86652	Hs.178294	ESTs	27.00
	125535	NM_013243	Hs.22215	secretogranin III	23.80
	125646	AA628962	Hs.75209	protein kinase (cAMP-dependent, catalyti	23.20
	125684	AW589427	Hs.158849	Homo sapiens cDNA: FLJ21663 fis, clone C	21.20
55	125724	AL360190	Hs.285978	Homo sapiens mRNA full length insert cDN	48.80
	125847	AW161885	Hs.249034	ESTs	31.00
	125934	AA193325	Hs.32646	hypothetical protein FLJ21901	21.20
	126077	M78772	Hs.210836	ESTs	49.80
	126299	AW979155	Hs.298275	amino acid transporter 2	21.80
60	126395	AA68004	Hs.278956	hypothetical protein FLJ12929	71.00
	126433	AA326606		gb:EST28707 Cerebellum II Homo sapiens c	23.20
	126509	R47400	Hs.23850	ESTs	23.80
	126538	AB030656	Hs.17377	coronin, actin-binding protein, 1C	23.10
	126666	AA648885	Hs.151999	ESTs	36.00
65	126812	AB037860	Hs.173933	nuclear factor I/A	20.80
	126872	AW450979		gb:U1-H-B13-ela-a-12-0-ULs1 NCL_CGAP_Su	46.29
	127046	AA321948	Hs.293968	ESTs	22.80
	127431	AW771958	Hs.175437	ESTs, Moderately similar to PC4259 ferri	30.00
	127489	AA650250	Hs.272076	ESTs	20.80
70	127521	AW297206	Hs.164018	ESTs	25.20
	127742	AW293496	Hs.180138	ESTs	28.00
	127925	AA805151	Hs.3628	mitogen-activated protein kinase kinase	21.20
	127930	AA809672	Hs.123304	ESTs	20.54
	127966	AA830201	Hs.124347	ESTs	28.20
75	127987	AK022103	Hs.124511	ESTs	19.60
	128116	H07103	Hs.286014	Homo sapiens, clone IMAGE:3867243, mRNA	20.40
	128609	NM_003616	Hs.102456	survival of motor neuron protein interac	34.40
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2	53.80
	128949	AA009647	Hs.8850	a disintegrin and metalloproteinase doma	23.00
80	129168	AI132988	Hs.109052	chromosome 14 open reading frame 2	37.60
	129404	AI267700	Hs.317584	ESTs	28.60
	129527	AA769221	Hs.270847	delta-tubulin	40.80
	129574	AA026815	Hs.11463	UMP-CMP kinase	31.20
	129598	N30436	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	29.60
85	129785	H19006	Hs.184780	ESTs	72.20
	129970	AV655806	Hs.296198	chromosome 12 open reading frame 4	22.20

	130149	AW067805	Hs.172665	methyltetrahydrofolate dehydrogenase	29.60
	130199	Z48579	Hs.172028	a disintegrin and metalloproteinase domain	27.60
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	28.36
5	130466	W19744	Hs.180059	Homo sapiens cDNA FLJ20653 fis, clone KA	20.20
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	22.40
	130617	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	19.60
	130703	R77776	Hs.18103	ESTs	19.40
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	21.40
10	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	110.00
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	25.20
	131086	AL035461	Hs.2281	chromogranin B (secretogranin 1)	40.60
	131284	NM_001429	Hs.25272	E1A binding protein p300	24.60
	131775	AB014548	Hs.31921	KIAA0648 protein	21.00
15	131860	BE383676	Hs.334	Rho guanine nucleotide exchange factor (	33.40
	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	60.80
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	20.40
	132084	NM_002267	Hs.3886	karyopherin alpha 3 (importin alpha 4)	29.40
	132389	AA310393	Hs.190044	ESTs	32.40
20	132437	AA152108	Hs.4859	cyclin L alpha-6a	27.40
	132550	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	75.60
	132617	AF037335	Hs.5338	carbonic anhydrase XII	31.36
	132632	AI076916	Hs.5398	guanine monophosphate synthetase	32.40
	132672	W27721	Hs.54697	Cdc42 guanine exchange factor (GEF) 9	23.40
25	132742	AA025480	Hs.292812	ESTs, Weakly similar to T33468 hypotheti	61.20
	132771	Y10275	Hs.56407	phosphoserine phosphatase	22.33
	133070	U92649	Hs.64311	a disintegrin and metalloproteinase domain	23.50
	133153	AF070592	Hs.66170	HSKM-B protein	30.00
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	23.80
30	133282	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	51.60
	133350	AI499220	Hs.71573	hypothetical protein FLJ10074	33.00
	133592	AV652066	Hs.75113	general transcription factor IIA	82.00
	133658	AA319146	Hs.75426	secretogranin II (chromogranin C)	
	133865	AB011155	Hs.170290	discs, large (Drosophila) homolog 6	69.33
35	134032	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	33.20
	134125	NM_014781	Hs.50421	KIAA0203 gene product	31.60
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	30.60
	134321	BE538082	Hs.8172	ESTs, Moderately similar to A45010 X-in	23.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	49.20
40	134570	U66615	Hs.172280	SWI/SNF related, matrix associated, acti	20.20
	134753	NM_006482	Hs.173135	dual-specificity tyrosine-(Y)-phosphoryl	20.80
	135002	AA448542	Hs.251677	G antigen 7B	37.60
	135029	H58818	Hs.187579	hydroxysteroid (17-beta) dehydrogenase	53.40
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	31.60
45	135345	X53655	Hs.99171	neurotrophin 3	28.80

TABLE 4B shows the accession numbers for those primekeys lacking unigenes/Ds for Table 4A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
123619	371681_1	AA602964 AA609200
126433	127143_1	AA325606 AA099517 N89423
126872	142696_1	AW450979 AA136553 AA136656 AW419381 AA984358 AA492073 BE168945 AA809054 AW238038 BE011212 BE011359 BE011367 BE011368 BE011362 BE011215 BE011365 BE011363
106851	322947_1	AA458623 AA639708 AA485409 R22065 AA485570
118720	genbank_N73515	N73515
120515	genbank_AA258356	AA258356
117099	321871_1	H93699 H97976 H80036
101447	entrez_M21305	M21305
123130	genbank_AA487200	AA487200

Table 5A shows 680 genes up-regulated in squamous cell carcinoma or adenocarcinoma lung tumors relative to normal lung and chronically diseased lung. These genes were selected from 59680 probesets on the Eos/Altymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number							
	ExAccn:	Exemplar Accession number, Genbank accession number							
	UnigeneID:	Unigene number							
	Unigene Title:	Unigene gene title							
10	R1:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R2:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R3:	80th percentile of AI squamous cell carcinoma lung tumor samples divided by the 90th percentile of AI for normal and chronically diseased lung samples.							
	R4:	80th percentile of AI adenocarcinoma lung tumor samples divided by the 80th percentile of AI for squamous cell carcinoma lung tumor samples.							
15	R5:	70th percentile of AI for squamous cell carcinoma and adenocarcinoma lung tumor samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples divided by 90th percentile of AI for normal and chronically diseased lung samples minus the 15th percentile of AI for all normal lung, chronically diseased lung and tumor samples							
20	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2	R3	R4	R5
	100035			AFFX control: GAPDH					6.76
	100036			AFFX control: GAPDH					5.77
	100037			AFFX control: GAPDH					5.75
25	100071	A28102		Human GABAA receptor alpha-3 subunit		8.00			
	100114	X02308	Hs.82962	thymidylate synthetase					5.71
	100154	H50720	Hs.81892	KIAA0101 gene product	3.84				
	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	3.33				
	100188	AW247090	Hs.57101	minichromosome maintenance deficient (S.					4.52
	100202	BE294407	Hs.99910	phosphofructokinase, platelet					5.49
30	100216	AA489908	Hs.1390	proteasome (prosome, macropain) subunit,					5.67
	100269	NM_001949	Hs.1189	E2F transcription factor 3	2.55				
	100287	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e					5.66
	100297	AU077258	Hs.182429	protein disulfide isomerase-related prot					3.81
	100330	AW410976	Hs.77152	minichromosome maintenance deficient (S.					4.50
35	100335	AW247529	Hs.6793	platelet-activating factor acetylhydrola	5.07				
	100360	W70171	Hs.75939	uridine monophosphate kinase					4.82
	100372	NM_014791	Hs.184339	KIAA0175 gene product					3.79
	100474	NM_000699	Hs.300280	amylase, alpha 2A; pancreatic				15.65	
40	100486	T19006	Hs.10842	RAN, member RAS oncogene family					5.49
	100491	D56165	Hs.275163	non-metastatic cells 2, protein (NM23B)					4.17
	100516	D90278	Hs.11	carcinoembryonic antigen-related cell ad		7.20			
	100522	X51501	Hs.99949	prolactin-induced protein				14.20	
	100559	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermoly	3.10				
45	100576	X00356	Hs.37058	calitonin/calcitonin-related polypeptid				9.30	
	100629	AA015693	Hs.21291	mitogen-activated protein kinase kinase				20.60	
	100661	BE623001	Hs.132748	Homo sapiens ribosomal protein L39 mRNA,	3.85				
	100677	AA353686	Hs.57813	zinc ribbon domain containing, 1		8.60			
	100696	D14887	Hs.121686	general transcription factor IIA, 1 (37k				10.00	
	100709	N26539	Hs.100469	myeloid/lymphoid or mixed-lineage leukem			24.80		
50	100781	BE208491	Hs.295112	KIAA0618 gene product		7.60			
	100830	AC004770	Hs.4756	flap structure-specific endonuclease 1					7.99
	100867	U14622		gb:Human transketolase-like protein gene		10.20			
	100902	M16029	Hs.287270	ret proto-oncogene (multiple endocrine n		8.00			
55	100906	AU076916	Hs.5398	guanine monophosphate synthetase					5.16
	100960	J00124	Hs.117729	keratin 14 (epidermolysis bullosa simple	2.57				
	101045	J05614		gb:Human proliferating cell nuclear anti					4.69
	101061	NM_000175	Hs.180532	glucose phosphate isomerase					4.19
	101071	L02840	Hs.84244	potassium voltage-gated channel, Shab-re		12.91			
60	101124	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	3.12				
	101175	U82671	Hs.36980	melanoma antigen, family A, 2	3.50				
	101181	BE262621	Hs.73798	macrophage migration inhibitory factor (					5.69
	101204	L24203	Hs.82237	ataxia-telangiectasia group D-associated	4.08				
	101210	L29301	Hs.2353	opioid receptor, mu 1			6.40		
65	101216	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	2.53				
	101228	AA333387	Hs.82916	chaperonin containing TCP1, subunit 6A (					7.90
	101233	AL135173	Hs.878	sorbitol dehydrogenase					4.45
	101273	Z11933	Hs.182505	POU domain, class 3, transcription facto	8.50				
	101342	U52112	Hs.182018	interleukin-1 receptor-associated kinase					4.17
70	101346	A738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N				21.89	
	101369	NM_000892	Hs.1901	kallikrein B, plasma (Fletcher factor) 1				12.80	
	101396	BE267931	Hs.78996	proliferating cell nuclear antigen	3.24				
	101431	BE185289	Hs.1076	small proline-rich protein 1B (comitin)					7.90
	101448	NM_000424	Hs.195850	keratin 5 (epidermolysis bullosa simplex	8.31				
75	101462	AL035668	Hs.73853	bone morphogenetic protein 2				38.80	
	101466	BE262680	Hs.170197	glutamic-oxaloacetic transaminase 2, mit					4.01
	101484	AA053486	Hs.20315	interferon-induced protein with tetratri				12.00	
	101502	M26958		gb:Human parathyroid hormone-related pro	10.50				
	101505	AA307680	Hs.75692	asparagine synthetase					4.46
80	101526	NM_002197	Hs.154721	aconitase 1, soluble	4.02				
	101535	X57152	Hs.99853	fibrillarin					4.65
	101577	M34363	Hs.1041	v-ras avian UR2 sarcoma virus oncogene h				9.09	
	101649	AW959908	Hs.1690	heparin-binding growth factor binding pr	54.00				
	101663	NM_003528	Hs.2178	H2B histone family, member Q	5.59				
	101664	AA436989	Hs.121017	H2A histone family, member A	7.00				
85	101669	L24498	Hs.80409	growth arrest and DNA-damage-inducible,		7.60			

5	101695	M69136	Hs.135626	chymase 1, mast cell	4.79			
	101724	L11690	Hs.620	bulbous pemphigoid antigen 1 (230/240kD)	15.21			
	101748	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen)	55.50			
	101759	M80244	Hs.184601	solute carrier family 7 (cationic amino)			18.57	4.10
	101771	NM_002432	Hs.153837	myeloid cell nuclear differentiation ant				
10	101804	M86699	Hs.169840	TTK protein kinase	4.50			
	101809	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	140.00			
	101833	AU076442	Hs.117938	collagen, type XVII, alpha 1	2.56			
	101842	M93221	Hs.75182	mannose receptor, C type 1			12.80	
	101851	BE260964	Hs.82045	midkine (neurite growth-promoting factor				5.88
15	102002	NM_002484	Hs.81469	nucleoside binding protein 1 (E.coli Min		7.80		
	102039	AL134223	Hs.306099	aldo-keto reductase family 1, member C1				4.35
	102072	U09410	Hs.78743	zinc finger protein 131 (clone pHZ-10)		7.40		
	102083	T35901	Hs.75117	interleukin enhancer binding factor 2, 4				5.12
	102111	L36196	Hs.81884	sulfotransferase family, cytosolic, 2A,			12.00	
20	102123	NM_001809	Hs.1594	centromere protein A (17kD)	6.20			
	102154	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kalinin	2.62			
	102193	AL036335	Hs.313	secreted phosphoprotein 1 (osteopontin,	5.85			
	102217	AA829978	Hs.301613	JTV1 gene				6.18
	102224	NM_002810	Hs.148495	proteasome (prosome, macropain) 26S subu				4.49
25	102234	AW163390	Hs.278554	heterochromatin-like protein 1				5.80
	102251	NM_004398	Hs.41706	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	4.50			
	102305	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)				5.15
	102330	BE298063	Hs.77254	chromobox homolog 1 (Drosophila HP1 beta				4.17
	102340	U37055	Hs.278557	macrophage stimulating 1 (hepatocyte gro			9.33	
30	102348	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	8.87			
	102358	U39817	Hs.36820	Bloom syndrome	15.91			
	102394	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma		19.20		
	102404	NM_005429	Hs.79141	vascular endothelial growth factor C			14.00	
	102537	U57094	Hs.50477	RAB27A, member RAS oncogene family			12.00	
35	102581	AU077228	Hs.77258	enhancer of zeste (Drosophila) homolog 2				4.57
	102605	AJ435128	Hs.181369	ubiquitin fusion degradation 1-like				3.98
	102610	U65011	Hs.30743	preferentially expressed antigen in meta	77.50			
	102623	AW249285	Hs.37110	melanoma antigen, family A, 9	12.50			
	102642	AA205847	Hs.23016	G protein-coupled receptor		22.00		
40	102654	AV649889	Hs.24385	Human hbc647 mRNA sequence		12.00		
	102659	BE245169	Hs.211610	CUG triplet repeat, RNA-binding protein			12.80	
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	6.50			
	102672	U72066	Hs.29287	retinoblastoma-binding protein 8	8.50			
	102687	NM_007019	Hs.93002	ubiquitin carrier protein E2-C				9.24
45	102696	BE540274	Hs.239	forkhead box M1				5.54
	102758	U82321		gb:Homo sapiens clone 14.9B mRNA sequenc	6.60			
	102781	BE258778	Hs.108809	chaperonin containing TCP1, subunit 7 (e				3.78
	102784	U85658	Hs.61798	transcription factor AP-2 gamma (activat				4.26
	102824	U90916	Hs.82845	Homo sapiens cDNA: FLJ21930 fis, clone H		14.40		
50	102829	NM_006183	Hs.80962	neurotensin	8.00			
	102888	AJ346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1				5.50
	102892	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin		6.70		
	102913	NM_002275	Hs.80342	keratin 15	4.64			
	102935	BE561850	Hs.80506	small nuclear ribonucleoprotein polypept	2.93			
55	102951	X15218	Hs.2959	v-eki avian sarcoma viral oncogene homol			11.40	
	102983	BE397202	Hs.118638	non-metastatic cells 1, protein (NM23A)				7.26
	103023	AW500470	Hs.117950	multifunctional polypeptide similar to S	3.01			
	103036	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	27.90			
	103038	AA926960	Hs.334883	CDC28 protein kinase 1				8.79
60	103060	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin				4.27
	103099	AJ693251	Hs.8248	NADH dehydrogenase (ubiquinone) Fe-S pro	9.80			
	103119	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	4.05			
	103168	X53463	Hs.2704	glutathione peroxidase 2 (gastric/intestin	3.07			
	103185	NM_006825	Hs.74368	transmembrane protein (63kD), endoplasmic				5.62
65	103192	M22440	Hs.170009	transforming growth factor, alpha		7.40		
	103223	BE275607	Hs.1709	chaperonin containing TCP1, subunit 3 (g				4.70
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o		100.00		
	103316	X83301	Hs.324728	SMA5			9.80	
	103375	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	9.71			
70	103376	AL036166	Hs.323378	coated vesicle membrane protein	14.00			
	103385	NM_007069	Hs.37189	similar to rat HREV107			11.00	
	103391	X94453	Hs.114366	pyrroline-5-carboxylate synthetase (glut	2.93			
	103404	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,				5.15
	103430	BE564090	Hs.20716	translocase of inner mitochondrial membr				3.98
75	103446	X98834	Hs.79971	sal (Drosophila)-like 2			21.40	
	103476	Y07701	Hs.293007	aminopeptidase puromycin sensitive	13.00			
	103477	AJ011812	Hs.119018	transcription factor NRF		6.40		
	103478	BE514982	Hs.38991	S100 calcium-binding protein A2	5.02			
	103515	Y10275	Hs.56407	phosphoserine phosphatase	10.50			
80	103558	BE616547	Hs.2785	keratin 17	6.41			
	103580	AA328046	Hs.46405	polymerase (RNA) II (DNA directed) polyp				3.84
	103587	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	78.50			
	103594	AJ368680	Hs.816	SRY (sex determining region Y)-box 2	6.51			
	103636	NM_006235	Hs.2407	POU domain, class 2, associating factor	3.50			
85	103768	AF066009		gb:Homo sapiens full length insert cDNA				4.48
	103841	AA314821	Hs.38178	hypothetical protein FLJ23468	8.00			
	103847	AF219946	Hs.102237	tubby super-family protein	10.40			
	103913	AW967500	Hs.133543	ESTs			15.60	
	104094	AA418187	Hs.330515	ESTs		6.60		

	104150	AL122044	Hs.331633	hypothetical protein DKFZp566N034		26.00	
	104257	BE560621	Hs.9222	estrogen receptor binding site associate	6.80		
	104261	AW248364	Hs.5409	RNA polymerase I subunit			3.98
5	104331	AB040450	Hs.279862	cdk inhibitor p21 binding protein	6.80		
	104415	BE410992	Hs.258730	heme-regulated initiation factor 2-alpha	10.29		
	104558	R56678	Hs.88959	hypothetical protein MGC4816	4.21		
	104590	AW373062	Hs.83623	nuclear receptor subfamily 1, group I, m		15.79	
	104658	AA360954	Hs.27268	Homo sapiens cDNA: FLJ21933 fis, clone H		17.40	
10	104660	BE298565	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	6.40		
	104689	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr			6.55
	104754	AI206234	Hs.155924	cAMP responsive element modulator		10.00	
	104758	BE560269	Hs.7010	NPD002 protein			4.47
	104971	BE311926	Hs.15830	hypothetical protein FLJ12691	2.87		
15	105011	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	3.83		
	105012	AF098158	Hs.9329	chromosome 20 open reading frame 1	2.86		
	105026	AA809485	Hs.124219	hypothetical protein FLJ12934	11.00		
	105076	AI598252	Hs.37810	hypothetical protein MGC14833			5.01
	105132	AA148164	Hs.247280	HBV associated factor			3.99
20	105143	AI368836	Hs.24808	ESTs, Weakly similar to I38022 hypotheti		11.00	
	105158	AW976357	Hs.234545	hypothetical protein NUF2R	16.00		
	105176	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	4.32		
	105200	AA328102	Hs.24641	cytoskeleton associated protein 2	3.00		
	105264	AA227934		gbzr57e08.s1 Soares_NhHMPu_S1 Homo sapi		10.00	
25	105298	BE387790	Hs.26369	hypothetical protein FLJ20287	3.69		
	105409	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8		9.20	
	105460	AW296078	Hs.271271	Homo sapiens, clone IMAGE:4179986, mRNA,		7.80	
	105667	AA767526	Hs.22030	paired box gene 5 (B-cell lineage specif	4.12		
	105743	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	3.82		
30	105782	H09748	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro		27.00	
	105848	AW954064	Hs.24951	ESTs		7.60	
	105891	U65984	Hs.289088	heat shock 90kD protein 1, alpha			4.14
	106019	AF221993	Hs.46743	McKusick-Kaufman syndrome		16.80	
	106069	BE568623	Hs.29899	ESTs, Weakly similar to G02075 transcrip		23.40	
35	106073	AL157441	Hs.17834	downstream neighbor of SON	9.50		
	106126	AA576953	Hs.22972	hypothetical protein FLJ13352	6.00		
	106159	AK001301	Hs.3487	hypothetical protein FLJ10439			3.95
	106220	D61329	Hs.32196	mitochondrial ribosomal protein L36			6.04
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S		13.20	
40	106300	Y10043	Hs.19114	high-mobility group (nonhistone chromoso			5.02
	106307	AA436174	Hs.37751	ESTs, Weakly similar to putative p150 [	6.60		
	106318	AA025610	Hs.9605	cleavage and polyadenylation specific fa			5.04
	106341	AF191020	Hs.5243	hypothetical protein, estradiol-induced			7.25
	106440	AA448563	Hs.151393	glutamate-cysteine ligase, catalytic sub		13.80	
45	106481	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	4.75		
	106586	AA243637	Hs.57787	ESTs		10.84	
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr		45.60	
	106654	AW075485	Hs.286049	phosphoserine aminotransferase	28.00		
	106785	Y15227	Hs.20149	deleted in lymphocytic leukemia, 1	3.00		
50	106813	C05766	Hs.181022	CGI-07 protein		11.40	
	106895	AK001826	Hs.25245	hypothetical protein FLJ11269		6.00	
	106913	AI219346	Hs.86178	M-phase phosphoprotein 9	6.56		
	106919	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S			4.27
	107054	AI076459	Hs.15978	KIAA1272 protein		34.80	
55	107059	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E. coli Re	4.71		
	107098	AI823593	Hs.27688	ESTs		24.80	
	107104	AU076640	Hs.15243	nuclear protein 1 (120kD)			7.05
	107129	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.60		
	107198	AV657225	Hs.9846	KIAA1040 protein		19.20	
60	107203	D20426	Hs.41639	programmed cell death 2		7.60	
	107217	AL080235	Hs.35861	DKFZP586E1621 protein	9.50		
	107284	NM_006829	Hs.187958	solute carrier family 6 (neurotransmitte	2.71		
	107318	T74445	Hs.5957	Homo sapiens clone 24416 mRNA sequence		8.71	
	107516	X57152	Hs.99853	fibrillarin			4.33
65	107529	BE515065	Hs.296585	nuclear protein (KKE/D repeat)			4.00
	107728	AA019551	Hs.294151	Homo sapiens, clone IMAGE:3603838, mRNA,	10.80		
	107851	AA022953	Hs.61172	EST		8.00	
	107901	L42612	Hs.335952	keratin 6B	3.40		
	107922	BE153855	Hs.61480	Ig superfamily receptor LNIR	2.88		
70	107932	AW392555	Hs.18878	hypothetical protein FLJ21620	7.50		
	108015	AW298357	Hs.49927	protein kinase NYD-SP15		23.40	
	108056	AA043675	Hs.52633	ESTs		12.80	
	108075	AI867370	Hs.139709	hypothetical protein FLJ12572		12.80	
	108187	BE245374	Hs.27842	hypothetical protein FLJ11210	7.00		
75	108296	N31256	Hs.161623	ESTs	6.60		
	108305	AA071391		gbzm61e06.r1 Stratagene fibroblast (937		11.80	
	108393	AA075211		gbzm86a08.r1 Stratagene ovarian cancer		11.80	
	108480	AL133092	Hs.68055	hypothetical protein DKFZp4340428		20.80	
	108554	AA084948		gbzm13b09.s1 Stratagene hNT neuron (937	6.40		
80	108573	AA086005		gbzm84c04.s1 Stratagene colon (937204)		25.40	
	108584	AA088326	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	9.60		
	108597	AK000292	Hs.278732	hypothetical protein FLJ20285		14.60	
	108695	AB029000	Hs.70823	KIAA1077 protein	3.00		
	108699	AA121514	Hs.70832	ESTs		10.00	
85	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c		11.00	
	108780	AU076442	Hs.117938	collagen, type XVII, alpha 1	11.21		



	108810	AW295847	Hs.71331	hypothetical protein MGC5350	8.50			
	108816	AA130884	Hs.270501	ESTs, Moderately similar to ALU2_HUMAN		7.40		
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	4.00			
5	108860	AA133334	Hs.128911	ESTs	6.09			
	108937	AL050107	Hs.24341	transcriptional co-activator with PDZ-bl	3.00			
	109010	NM_007240	Hs.44229	dual specificity phosphatase 12	2.69			
	109121	BE389367	Hs.49767	NADH dehydrogenase (ubiquinone) Fe-S pro			4.53	
	109166	AA219691	Hs.73625	RAB6 interacting, kinasin-like (rabkines	10.58			
10	109227	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21		9.00		
	109415	U80736	Hs.110826	trinucleotide repeat containing 9		51.40		
	109418	AB66946	Hs.161707	ESTs			11.00	
	109454	AA232255	Hs.295232	ESTs, Moderately similar to A46010 X-li		17.60		
	109502	AW967069	Hs.211556	hypothetical protein MGC5487		9.49		
15	109543	AA564994	Hs.222851	ESTs	12.67			
	109548	H17800	Hs.7154	ESTs			10.40	
	109580	AB037734	Hs.4993	KIAA1313 protein		33.20		
	109700	F09609		gb:HSC33H092 normalized infant brain cDN			16.00	
	109704	AF438880	Hs.12876	ESTs		11.00		
20	109792	R49625		gb:yg51f03.s1 Soares infant brain 1NIB H			12.60	
	109981	BE546208	Hs.26090	hypothetical protein FLJ20272	4.00			
	109998	AL042201	Hs.21273	transcription factor NYD-sp10		7.80		
	110039	H11938	Hs.21907	histone acetyltransferase		7.00		
	110156	AA581322	Hs.4213	hypothetical protein MGC16207			4.24	
25	110500	AA907723	Hs.36962	ESTs	4.50			
	110551	AW450381	Hs.14529	ESTs		8.60		
	110561	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	3.06			
	110854	BE612992	Hs.27931	hypothetical protein FLJ10607 similar to		6.80	8.80	
	110886	AW274992	Hs.72249	three-PDZ containing protein similar to				
30	110916	BE178102	Hs.24349	ESTs		6.80		
	111003	N52980	Hs.83765	dihydrofolate reductase			16.80	
	111337	AA837396	Hs.263925	LIS1-interacting protein NUDE1, rat homo	2.54			
	111434	R01608	Hs.142736	ESTs			9.80	
	111439	AF476429	Hs.19238	ESTs			10.40	
35	111540	U82670	Hs.9786	zinc finger protein 275		15.40		
	111597	R11499	Hs.189716	ESTs			9.20	
	111895	T80581	Hs.12723	Homo sapiens clone 25153 mRNA sequence		6.80		
	111929	AF027208	Hs.112360	prominin (mouse)-like 1			14.67	
	112054	R43590		gb:yc65g02.s1 Soares infant brain 1NIB H	10.80			
40	112210	R49645	Hs.7004	ESTs			10.20	
	112244	AB029000	Hs.70823	KIAA1077 protein	2.99			
	112382	R59904		gb:yh07g12.s1 Soares infant brain 1NIB H		6.60		
	112392	R60763	Hs.193274	ESTs, Moderately similar to I57588 HStr		7.10		
	112442	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	3.00			
45	112539	R70318	Hs.339730	ESTs			37.20	
	112772	AI992283	Hs.35437	ESTs, Moderately similar to I38026 MLN 6			14.60	
	112869	BE261750	Hs.4747	dyskeratosis congenita 1, dyskerin				4.83
	112935	R71449	Hs.268760	ESTs	2.73			
	112970	AA694010	Hs.6932	Homo sapiens clone 23809 mRNA sequence			12.00	
50	112973	AB033023	Hs.318127	hypothetical protein FLJ10201	11.50			
	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f		10.89		
	113063	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	15.00		15.31	
	113073	N39342	Hs.103042	microtubule-associated protein 1B				
	113078	T40444	Hs.118354	CAT56 protein	7.00			
55	113238	R45467	Hs.189813	ESTs			41.20	
	113591	T91881	Hs.200597	KIAA0563 gene product			9.40	
	113702	T97307		gb:ye53h05.s1 Soares fetal liver spleen	25.00			
	113844	AI369275	Hs.243010	Homo sapiens cDNA FLJ14445 fis, clone HE			13.91	
	113984	R96696	Hs.35598	ESTs		7.80		
60	114073	R44953	Hs.22908	Homo sapiens mRNA; cDNA DKFZp434J1027 (f		7.20		
	114162	AF155681	Hs.22265	pyruvate dehydrogenase phosphatase	3.42		6.74	
	114208	AL049466	Hs.7859	ESTs				
	114251	H15261	Hs.21948	ESTs			33.20	
	114285	R44338	Hs.22974	ESTs			13.20	
65	114313	H18456	Hs.27946	ESTs			10.00	
	114339	AA782845	Hs.22790	ESTs		7.80		
	114407	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f				4.14
	114560	AI452469	Hs.165221	ESTs			9.80	
	114699	AA127386		gb:zn90d09.r1 Stratagene lung carcinoma		7.60		
70	114767	AI859865	Hs.154443	minichromosome maintenance deficient (S	3.21			
	114793	AA158245		gb:zo76c03.s1 Stratagene pancreas (93720		6.00		
	114833	AI417215	Hs.87159	hypothetical protein FLJ12577			11.40	
	115047	BE270930	Hs.82916	chaperonin containing TCP1, subunit 6A (				4.31
	115060	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3				4.03
75	115097	AA256213	Hs.72010	ESTs			35.40	
	115113	AA256460		gb:zr81a04.s1 Soares_NhHMPu_S1 Homo sapi			15.20	
	115123	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m				4.19
	115134	AW968073	Hs.194331	ESTs, Highly similar to A56713 inositol			12.40	
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	25.00			
80	115347	AA356792	Hs.334824	hypothetical protein FLJ14825		7.00		
	115414	AA662240	Hs.283099	AF15q14 protein	3.25			
	115522	BE614397	Hs.333893	c-Myc target JPO1	3.68			
	115536	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	10.50			
	115566	AI142336	Hs.43977	Human DNA sequence from clone RP11-196N1			24.40	
85	115645	AI207410	Hs.69280	Homo sapiens, clone IMAGE:3636299, mRNA,	4.17			
	115648	AW016811	Hs.234478	Homo sapiens cDNA: FLJ22648 fis, clone H		6.00		

	115652	BE093589	Hs.38178	hypothetical protein FLJ23468	3.81			
	115697	D31382	Hs.63325	transmembrane protease, serine 4	62.14			
	115793	AA424883	Hs.70333	hypothetical protein MGC10753			11.80	
	115816	BE042915	Hs.287588	Homo sapiens cDNA FLJ13675 fis, clone PL			9.71	
5	115892	AA291377	Hs.50831	ESTs		27.40		
	115905	AI767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	2.53			
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAPI_HUMAN DEATH	11.82			
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970			34.29	
	115978	AL035864	Hs.69517	cDNA for differentially expressed CO16 g			8.23	
10	115985	AA447709	Hs.268115	ESTs, Weakly similar to T08599 probable	3.00			
	116090	AI591147	Hs.61232	ESTs	5.17			
	116096	AA682382	Hs.59982	ESTs		8.20		
	116127	AF126743	Hs.279884	DNAJ domain-containing		10.60		
	116157	BE439838	Hs.44298	mitochondrial ribosomal protein S17			5.82	
15	116190	AI949095	Hs.67776	ESTs, Weakly similar to T22341 hypotheti			4.08	
	116278	NM_003686	Hs.47504	exonuclease 1	9.50			
	116335	AK001100	Hs.41690	desmocollin 3	3.67			
	116496	AW450694	Hs.21433	hypothetical protein DKFZp547J036	7.00			
20	116503	AI925316	Hs.212617	ESTs			12.60	
	116674	AI768015	Hs.92127	ESTs		32.00		
	116929	AA586922	Hs.80475	polymerase (RNA) II (DNA directed) polyp	7.60			
	116973	AI702054	Hs.166982	phosphatidylinositol glycan, class F	9.80			
	116993	AA117023	Hs.40478	ESTs			10.20	
	117079	H92325		gbys85f05.s1 Soares retina N2b4HR Homo			15.20	
25	117317	AI263517	Hs.43322	ESTs			13.40	
	117326	N23629	Hs.241420	Homo sapiens mRNA for KIAA1756 protein,			20.60	
	117396	W20128	Hs.296039	ESTs			10.60	
	117412	N32536	Hs.42645	ESTs			16.00	
	117519	N32528	Hs.146286	kinesin family member 13A			9.11	
30	117693	AW179019	Hs.112110	mitochondrial ribosomal protein L42			4.01	
	117721	N46100	Hs.93939	EST			19.80	
	117881	AF161470	Hs.260622	butyrate-induced transcript 1	2.71			
	117903	AA768283	Hs.47111	ESTs			17.80	
35	117992	AI015709	Hs.172089	Homo sapiens mRNA; cDNA DKFZp586I2022 (f			4.17	
	118013	AI674126	Hs.94031	ESTs			10.60	
	118017	AI813444	Hs.42197	ESTs		8.82		
	118186	N22886	Hs.42380	ESTs	7.00			
	118325	AI868065	Hs.166184	intersectin 2			13.80	
40	118367	N64269	Hs.48946	EST		6.14		
	118368	N64339	Hs.48956	gap junction protein, beta 6 (connexin 3	3.14			
	118472	AL157545	Hs.42179	bromodomain and PHD finger containing, 3		12.40		
	118709	AA232970	Hs.293774	ESTs			12.20	
	119025	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	4.50			
45	119027	AF086161	Hs.114611	hypothetical protein FLJ11808	3.22			
	119052	R10889		gbv38d02.s1 Soares fetal liver spleen	9.60			
	119164	AF221993	Hs.46743	McKusick-Kaufman syndrome		6.60		
	119186	AI979147	Hs.101265	hypothetical protein FLJ22593			10.80	
	119243	T12603		gb:CHR0123 Chromosome 9 exon II Homo sa			9.44	
50	119490	AA195276	Hs.263858	ESTs, Moderately similar to B34087 hypot			11.80	
	119499	AI918906	Hs.55080	ESTs			14.80	
	119599	W45552		gbzcc26d03.s1 Soares_senescent_fibroblas	12.60			
	119780	NM_016625	Hs.191381	hypothetical protein	17.00			
	119845	W79123	Hs.58561	G protein-coupled receptor 87	13.50			
55	119941	AA689485	Hs.58896	ESTs	8.00			
	119994	AA642402	Hs.59142	ESTs	7.73			
	120102	W67353	Hs.170218	KIAA0251 protein		39.60		
	120104	AK000123	Hs.180479	hypothetical protein FLJ20116	2.91			
	120294	AK000059	Hs.153861	Homo sapiens NY-REN-62 antigen mRNA, par		8.20		
60	120486	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	8.73			
	120699	AA804448	Hs.104463	ESTs	7.00			
	120699	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos			10.00	
	120715	AA292700		gbzcs59a05.s1 NCI_CGAP_GCB1 Homo sapiens	9.40			
	120821	Y19062	Hs.96870	stauferin (Drosophila, RNA-binding protein			13.80	
65	120859	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	9.00			
	120880	AA360240	Hs.97019	EST	15.50			
	120983	AA398209	Hs.97587	EST		27.66		
	121034	AL389951	Hs.271623	nucleoporin 50kD		20.80		
	121121	AA399371	Hs.189095	similar to SALL1 (sal (Drosophila)-like	22.80			
70	121313	AA402713	Hs.97872	ESTs			10.00	
	121369	AW450737	Hs.128791	CGI-09 protein	25.71			
	121376	AA448103	Hs.187958	solute carrier family 6 (neurotransmitte			5.42	
	121476	AA412311	Hs.97903	ESTs	8.30			
	121609	AA868939	Hs.97888	ESTs	8.59			
	121553	AA412488	Hs.48820	TATA box binding protein (TBP)-associat	18.50			
75	121763	AK000552	Hs.323518	WD repeat domain 5	7.00			
	121838	AA425580	Hs.98441	ESTs			10.40	
	121857	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	6.00			
	121991	AA430058	Hs.98649	EST			12.20	
80	122089	AW016543	Hs.98682	hypothetical protein FKSG32		8.60		
	122105	AW241685	Hs.98699	ESTs		6.14		
	122163	AA435702	Hs.98829	EST			10.40	
	122318	AA429743		gbzv0b05.r1 Soares_testis_NHT Homo sap			18.20	
	122335	AA443258	Hs.241551	chloride channel, calcium activated, fam	13.50			
	122338	AA443311	Hs.98998	ESTs	4.80			
85	122414	AI313473	Hs.99087	ESTs, Weakly similar to S47073 finger pr	8.00			

5	122512	AF053305	Hs.98658	budding uninhibited by benzimidazoles 1	8.80		
	122516	AA449352	Hs.99217	ESTs		9.40	
	122702	AI220089	Hs.99439	ESTs	9.20		
	122852	AI580056	Hs.98992	ESTs		10.40	
	122925	AW268962	Hs.111335	ESTs	6.80		
10	123005	AW369771	Hs.52620	integrin, beta 8		12.60	
	123044	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro			5.35
	123160	AA488687	Hs.284235	ESTs, Weakly similar to I38022 hypothe		6.06	
	123315	AA486369		gbzv37d10.s1 Soares ovary tumor NbHOT H		12.40	
	123329	Z47542	Hs.179312	small nuclear RNA activating complex, po		11.80	
15	123497	AA765256	Hs.135191	ESTs, Weakly similar to unnamed protein	12.00		
	123518	AL035414	Hs.21068	hypothetical protein		13.00	
	123519	AW015887	Hs.112574	ESTs	12.20		
	123614	AK000492	Hs.98806	hypothetical protein		7.80	
	123616	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L		10.60	
20	123673	BE550112	Hs.158549	ESTs, Weakly similar to T2D3_HUMAN TRANS	23.00		
	123727	AI083986	Hs.282977	hypothetical protein FLJ13490	7.00		
	123731	AA609839		gbas62f01.s1 Stratagene lung carcinoma		9.80	
	123752	AA227714	Hs.179703	KIAA0129 gene product	3.50		
	123900	AA621223	Hs.112953	EST		12.80	
25	124006	AI147155	Hs.270016	ESTs	97.00		
	124059	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypothe	3.02		
	124069	AF134160	Hs.7327	claudin 1		27.80	
	124191	T96509	Hs.248549	ESTs, Moderately similar to S65657 alpha		35.80	
	124273	AA457211	Hs.8858	bromodomain adjacent to zinc finger doma	7.20		
30	124297	AL080215	Hs.102301	Homo sapiens mRNA; cDNA DKFZp586j0323 (f		11.00	
	124305	AW953221		gb:EST375294 MAGE resequences, MAGH Homo		16.00	
	124676	AI360119.comp	Hs.181013	phosphoglycerate mutase 1 (brain)			6.08
	124674	BE550182	Hs.127826	RalGEF-like protein 3, mouse homolog		21.00	
	124904	AK000483	Hs.93872	KIAA1682 protein	9.40		
35	124969	AI650360	Hs.100256	ESTs		10.80	
	125000	T58615	Hs.110640	ESTs		9.80	
	125201	AA693960	Hs.103158	ESTs, Weakly similar to T33296 hypothe	7.60		
	125266	W90022	Hs.186809	ESTs, Highly similar to LCT2_HUMAN LEUKO	6.59		
	125299	T32882	Hs.102720	ESTs		9.57	
40	125356	AI057052	Hs.133554	ESTs, Weakly similar to Z195_HUMAN ZINC		14.00	
	125370	AA256743	Hs.134158	Homo sapiens, Similar to KIAA0092 gene p	8.20		
	125418	AA777690	Hs.188501	ESTs		13.20	
	125433	AL162066	Hs.54320	hypothetical protein DKFZp762D096	21.40		
	125437	AI609449	Hs.140197	ESTs	6.96		
45	125446	BE219987	Hs.166982	phosphatidylinositol glycan, class F	8.80		
	125711	AA305800	Hs.5672	hypothetical protein AF140225		11.20	
	125756	BE174587	Hs.289721	growth arrest specific transcript 5			4.31
	125757	AI274906	Hs.166835	ESTs, Highly similar to 1814460A p53-ass		15.60	
	125769	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	3.20		
50	125839	AW836261	Hs.337717	ESTs	8.20		
	125850	W85858	Hs.99804	ESTs	2.65		
	125875	H14480		gbym18b09.r1 Soares infant brain 1NIB H	7.40		
	125924	BE272506	Hs.82109	syndecan 1			4.23
	125972	AI927475	Hs.35406	ESTs, Highly similar to unnamed protein		3.98	
55	126034	H50340		gbvyr39b04.r1 Soares fetal liver spleen		10.60	
	126327	AA432266	Hs.44648	ESTs	11.60		
	126345	N49713		gbvyr23f06.s1 Soares fetal liver spleen	6.67		
	126435	AW614529	Hs.285847	CGI-19 protein		10.60	
	126487	AA283809	Hs.184601	solute carrier family 7 (cationic amino			4.38
60	126521	AI475110	Hs.203933	ESTs	6.60		
	126522	W31912		gbzcv76d03.s1 Pancreatic Islet Homo sapi		14.80	
	126543	AL035864	Hs.69517	cDNA for differentially expressed CO16 g			4.01
	126567	AA058394	Hs.57887	ESTs, Weakly similar to KIAA0758 protein	7.80		
	126605	AA676910		gbzcv65h07.s1 Soares fetal liver spleen		11.60	
65	126627	AA497044	Hs.20887	hypothetical protein FLJ10392		14.60	
	126628	N49776	Hs.170994	hypothetical protein MGC10946	8.00		
	126737	AW976516	Hs.283707	Homo sapiens cDNA: FLJ21354 fis, clone C	2.92		
	126795	AW975076	Hs.172589	nuclear phosphoprotein similar to S. cer	7.50		
	126802	AW805510	Hs.97056	hypothetical protein FLJ21634	11.60		
70	126892	AF121855	Hs.284291	sorting nexin 6	3.50		
	126928	AA480902	Hs.137401	ESTs		22.83	
	126979	AA210954		gbzcv89h10.r1 Stratagene hNT neuron [937		11.80	
	126986	AI279892	Hs.46801	sorting nexin 14		11.60	
	126992	AI809521		gbvrf30e03.x1 Soares_NFL_T_GBC_S1 Homo s		20.80	
75	127066	R25066		gbvyp42c07.r1 Soares infant brain 1NIB H		27.60	
	127099	AA347668		gb:EST54026 Fetal heart II Homo sapiens		21.60	
	127139	AA830233	Hs.293585	ESTs		11.20	
	127209	AA305023	Hs.81964	SEC24 (S. cerevisiae) related gene famil	3.10		
	127221	BE062109	Hs.241551	chloride channel, calcium activated, fam	2.76		
80	127225	AA315933	Hs.120879	ESTs		16.80	
	127313	AK002014	Hs.47546	Homo sapiens cDNA FLJ11458 fis, clone HE	14.00		
	127444	AW978474	Hs.7560	Homo sapiens mRNA for KIAA1729 protein,		13.60	
	127500	AW971353	Hs.162115	ESTs	11.20		
	127524	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin	7.80		
85	127540	N45572	Hs.105362	Homo sapiens, clone MGC:18257, mRNA, com	3.53		
	127599	AA613204	Hs.150399	ESTs		13.80	
	127609	X80031	Hs.530	collagen, type IV, alpha 3 (Goodpasture		28.00	
	127662	W80755	Hs.8294	KIAA0196 gene product		19.80	
	127668	AI343257	Hs.139993	ESTs		11.20	

	127746	AI239495	Hs.120189	ESTs			14.18
	127812	AA741368	Hs.291434	ESTs	4.50		
	127817	AA836641	Hs.163085	ESTs			24.60
5	127959	AI302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L			9.20
	127960	AI613226	Hs.41569	phosphatidic acid phosphatase type 2A			16.83
	127969	F08498	Hs.93748	Homo sapiens cDNA FLJ14676 fis, clone NT	13.60		
	128015	Z21169	Hs.334659	hypothetical protein MGC14139	7.00		
	128027	AI433721	Hs.164153	ESTs			37.40
10	128077	AI310330	Hs.128720	ESTs			9.60
	128166	NM_006147	Hs.11801	interferon regulatory factor 6			9.24
	128226	AI284940	Hs.289082	GM2 ganglioside activator protein	19.00		
	128305	AI954968	Hs.278009	matrix Gla protein			10.40
	128341	AA191420	Hs.185030	ESTs	9.00		
15	128527	AA504583	Hs.101047	transcription factor 3 (E2A immunoglobul			4.30
	128539	R46163	Hs.258618	ESTs	12.60		
	128568	H12912	Hs.274691	adenylylate kinase 3			4.56
	128572	AA933022	Hs.256583	interleukin enhancer binding factor 3, 9		10.00	
	128777	AI878918	Hs.10526	cysteine and glycine-rich protein 2		16.80	
20	128781	N71826	Hs.105465	small nuclear ribonucleoprotein polypept			4.48
	128796	AJ000152	Hs.105924	defensin, beta 2	8.12		
	128920	AA622037	Hs.166468	programmed cell death 5			4.62
	128924	BE279383	Hs.26557	plakophilin 3			4.04
	128971	H05132	Hs.107510	ESTs	12.60		
25	129008	AL079648	Hs.301088	ESTs	8.80		
	129041	BE382756	Hs.169902	solute carrier family 2 (facilitated glu			6.05
	129075	BE250162	Hs.83765	dihydrofolate reductase	2.59		
	129105	AI769160	Hs.108681	Homo sapiens brain tumor associated prot		6.67	
	129189	AB023179	Hs.9059	KIAA0962 protein	8.00		
30	129229	AF013758	Hs.109543	polyadenylate binding protein-interactin	4.00		
	129241	AI878857	Hs.109706	hematological and neurological expressed			4.06
	129300	W94197	Hs.110165	ribosomal protein L26 homolog	2.55		
	129404	AI267700	Hs.317584	ESTs	18.00		
	129457	X61959	Hs.207776	aspartylglucosaminidase	6.50		
35	129466	L42583	Hs.334309	keratin 6A	12.94		
	129494	AI148976	Hs.112062	ESTs			11.00
	129605	AF061812	Hs.115947	keratin 16 (focal non-epidermolytic palm			4.46
	129641	AI911527	Hs.11805	ESTs			12.00
	129665	AW163331	Hs.118778	KDEL (Lys-Asp-Glu-Leu) endoplasmic retic			4.70
40	129703	BE388665	Hs.179999	Homo sapiens, clone IMAGE:3457003, mRNA			4.02
	129720	AA156214	Hs.12152	APMCF1 protein			5.71
	129748	M16707	Hs.123053	H4 histone, family 2	3.50		
	129890	AI868872	Hs.282804	hypothetical protein FLJ27704			4.21
	129896	BE295568	Hs.13225	UDP-Gal-beta-GlcNAc beta 1,4- galactosylt	2.56		
45	129945	BE514376	Hs.165998	PAI-1 mRNA-binding protein			4.03
	130010	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34		7.00	
	130026	T40480	Hs.332112	EST		6.40	
	130080	X14850	Hs.147097	H2A histone family, member X			4.65
50	130149	AW067805	Hs.172655	methyleneetetrahydrofolate dehydrogenase	2.74		
	130285	AA063546	Hs.75981	ubiquitin specific protease 14 (ubiquitin-gua		7.40	
	130441	U63630	Hs.155637	protein kinase, DNA-activated, catalytic			3.91
	130482	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	4.87		
	130500	AB007913	Hs.158291	KIAA0444 protein			9.80
55	130524	U89995	Hs.159234	forkhead box E1 (thyroid transcription f		13.40	
	130541	X05608	Hs.211584	neurofilament, light polypeptide (58kD)		8.20	
	130553	AF062649	Hs.252587	pituitary tumor-transforming 1			6.06
	130567	AA383092	Hs.1608	replication protein A3 (14kD)		7.00	
	130577	M69241	Hs.162	insulin-like growth factor binding prote	3.04		
60	130627	BE003064	Hs.1695	matrix metalloproteinase 12 (macrophage	3.87		
	130648	AI458165	Hs.17286	hypothetical protein MGC2376			16.20
	130697	L29472	Hs.1802	major histocompatibility complex, class			17.80
	130744	H59696	Hs.18747	POP7 (processing of precursor, S. cerevi			5.28
	130800	AI187292	Hs.19574	hypothetical protein MGC5469			4.43
	130867	NM_001072	Hs.284239	UDP glycosyltransferase 1 family, polype	16.84		
65	130869	J03626	Hs.2057	uridine monophosphate synthetase (orotat			4.92
	130925	AF093419	Hs.168378	multiple PDZ domain protein			9.60
	130994	W17044	Hs.327337	ESTs		12.40	
	131028	AI879165	Hs.2227	CCAAT/enhancer binding protein (C/EBP),	10.21		
70	131031	NM_001650	Hs.288650	aquaporin 4			9.80
	131041	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,			9.60
	131058	W28545	Hs.101514	hypothetical protein FLJ10342			17.00
	131090	AI143139	Hs.2288	visinin-like 1	2.74		
	131112	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f		8.80	
75	131148	AW953575	Hs.303125	p53-induced protein PIGPC1	3.12		
	131185	BE280074	Hs.23860	cyclin B1	3.07		
	131200	BE540516	Hs.293732	hypothetical protein MGC3195	3.07		
	131219	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	2.87		
	131257	AW339037	Hs.24908	ESTs			14.67
	131375	AW293165	Hs.143134	ESTs		19.20	
80	131460	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.50		
	131476	AI521663	Hs.334644	hypothetical protein FLJ14668	15.00		
	131510	BE245374	Hs.27842	hypothetical protein FLJ11210		7.80	
	131646	BE302464	Hs.30057	MRS2 (S. cerevisiae)-like, magnesium hom		7.00	
	131786	BE000971	Hs.306083	Novel human gene mapping to chromosome 22	2.65		
85	131839	AB014533	Hs.33010	KIAA0633 protein			35.20
	131843	AA192315	Hs.184062	putative Rab5-interacting protein			4.11

	131877	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	19.00		
	131885	BE502341	Hs.3402	ESTs	6.48		
	131921	AA456093	Hs.34720	ESTs		8.40	
5	131945	NM_002916	Hs.35120	replication factor C (activator 1) 4 (37	56.00		
	131958	NM_014082	Hs.3566	ART-4 protein			3.82
	131965	W79283	Hs.35962	ESTs	3.03		
	132000	AW247017	Hs.36978	melanoma antigen, family A, 3		9.80	
	132040	NM_001196	Hs.315689	Homo sapiens cDNA: FLJ22373 fis, clone H	3.30		
10	132109	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	21.00		
	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		8.40	
	132162	AA315805	Hs.94560	desmoglein 2			12.25
	132164	AI752235	Hs.41270	procollagen-lysine, 2-oxoglutarate 5-dio	2.70		
	132180	NM_004460	Hs.418	fibroblast activation protein, alpha	2.71		
15	132181	AW961231	Hs.16773	Homo sapiens clone TCCCA00427 mRNA sequ	3.83		
	132182	NM_014210	Hs.70499	ecotropic viral integration site 2A			13.20
	132231	AA662910	Hs.42635	hypothetical protein DKFZp434K2435	9.50		
	132277	AK001745	Hs.184628	hypothetical protein FLJ10883	4.50		
	132328	NM_014787	Hs.44896	DnaJ (Hsp40) homolog, subfamily B, membe			9.20
20	132394	AK001680	Hs.30488	DKFZP434F091 protein			19.80
	132424	AA417878	Hs.48401	ESTs, Moderately similar to ALU8_HUMAN A		8.60	
	132528	T78736	Hs.50758	SMC4 (structural maintenance of chromoso		27.40	
	132543	BE568452	Hs.5101	protein regulator of cytokinesis 1	4.38		
	132544	L19778	Hs.51011	H2A histone family, member P		7.00	
25	132560	AW969253	Hs.170195	bone morphogenetic protein 7 (osteogenic	2.64		
	132552	BE621985	Hs.296922	thiopurine S-methyltransferase			15.83
	132581	AK000631	Hs.52256	hypothetical protein FLJ20624		6.60	
	132617	AF037335	Hs.5338	carbonic anhydrase XII	4.95		
	132638	AI796870	Hs.54277	DNA segment on chromosome X (unique) 992		8.20	
	132653	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalfi	4.38		
30	132669	W38586	Hs.293981	guanine nucleotide binding protein (G pr			4.36
	132710	W74001	Hs.55279	serine (or cysteine) proteinase inhibito	4.60		
	132771	Y10275	Hs.56407	phosphoserine phosphatase	3.71		
	132799	W73311	Hs.169407	SAC2 (suppressor of actin mutations 2,			9.48
	132833	U78525	Hs.57783	eukaryotic translation initiation factor			5.83
35	132892	AW834050	Hs.9973	tensin			12.00
	132906	BE513337	Hs.234896	geminin	3.09		
	132959	AW014195	Hs.61472	ESTs, Weakly similar to YAE5_YEAST HYPOT			3.87
	132962	AA576635	Hs.6153	CGI-48 protein	3.50		
40	132990	X77343	Hs.334334	transcription factor AP-2 alpha (activat	6.18		
	132994	AA112748	Hs.279905	clone HQ0310 PRO0310p1	3.19		
	133000	ALD42444	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	2.96		
	133050	X73424	Hs.63788	propionyl Coenzyme A carboxylase, beta p	2.55		
	133083	BE244588	Hs.6456	chaperonin containing TCP1, subunit 2 (b			4.00
45	133086	L17131	Hs.139800	high-mobility group (nonhistone chromoso			8.96
	133134	AF198620	Hs.65648	RNA binding motif protein 8A			4.28
	133155	M58583	Hs.662	cerebellin 1 precursor		10.80	
	133181	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	3.00		
	133204	BE267696	Hs.254105	enolase 1, (alpha)			4.63
50	133412	U41493	Hs.73112	guanine nucleotide binding protein (G pr		12.50	
	133421	AF134160	Hs.7327	claudin 1	2.85		
	133451	AW970026	Hs.73818	ubiquinol-cytochrome c reductase hinge p			4.66
	133453	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept		6.80	
	133504	NM_004415	Hs.74316	desmoplakin (DPI, DPII)	6.14		
55	133506	BE562958	Hs.74346	hypothetical protein MGC14353			4.55
	133615	M62843	Hs.75236	ELAV (embryonic lethal, abnormal vision,			17.80
	133627	NM_002047	Hs.75280	glycyl-tRNA synthetase			4.85
	133649	U25849	Hs.75393	acid phosphatase 1, soluble			6.34
	133669	NM_006925	Hs.166975	splicing factor, arginine/serine-rich 5			14.00
60	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		6.11	
	133776	BE268649	Hs.177766	ADP-ribosyltransferase (NAD+; poly (ADP-			4.91
	133866	AB011155	Hs.170290	discs, large (Drosophila) homolog 5	3.07		
	133946	AJ001258	Hs.173879	NIPSNAP, C. elegans, homolog 1			4.60
	133973	N55540	Hs.78026	ESTs, Weakly similar to similar to ankyl			13.00
65	134047	BE262529	Hs.78771	phosphoglycerate kinase 1			3.85
	134098	BE513171	Hs.79086	mitochondrial ribosomal protein L3	2.56		
	134107	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte		8.20	
	134112	AW449809	Hs.79150	chaperonin containing TCP1, subunit 4 (d			4.08
	134158	U15174	Hs.79428	BCL2/adenovirus E1B 19kD-interacting pro	31.00		
70	134160	T98152	Hs.79432	fibrillin 2 (congenital contractural era		24.60	
	134168	AA398908	Hs.181634	Homo sapiens cDNA: FLJ23602 fis, clone L			6.71
	134185	AA285136	Hs.301914	neuronal specific transcription factor D			14.74
	134201	L35035	Hs.79886	ribose 5-phosphatase isomerase A (ribose 5		8.40	
	134272	X76040	Hs.278614	protease, serine, 15	4.50		
75	134276	BE083936	Hs.80976	antigen identified by monoclonal antibod		9.00	
	134353	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m			16.40
	134367	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	2.80		
	134380	AJ077143	Hs.179565	minichromosome maintenance deficient (S.	4.68		
	134423	H53497	Hs.83006	CGI-139 protein			3.84
	134469	AA279661	Hs.83753	small nuclear ribonucleoprotein polypept			5.81
80	134470	X54942	Hs.83758	CDC28 protein kinase 2			4.21
	134498	AW246273	Hs.84131	threonyl-tRNA synthetase			7.30
	134502	BE148534	Hs.84168	UV-B repressed sequence, HUR 7		13.60	
	134510	NM_002757	Hs.250870	mitogen-activated protein kinase kinase			9.70
	134548	N95406	Hs.333495	Deleted in split-hand/split-foot 1 regio			4.63
85	134654	AK001741	Hs.8739	hypothetical protein FLJ10879	6.00		

5	134724	AF045239	Hs.321576	ring finger protein 22		12.00
	134743	AA044163	Hs.89463	potassium large conductance calcium-acti	4.00	
	134781	AA374372	Hs.89626	parathyroid hormone-like hormone		25.20
	134806	AD001528	Hs.89718	spermine synthase		4.58
	134853	BE268326	Hs.90280	5-aminimidazole-4-carboxamide ribonucle		4.79
	134859	D26488	Hs.90315	KIAA0007 protein		6.20
	134891	RS1083	Hs.90787	ESTs		7.40
	134960	BE246400	Hs.285176	acetyl-Coenzyme A transporter	4.00	
10	134993	BE409809	Hs.301005	purine-rich element binding protein B		4.48
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	9.50	
	135080	AI761180	Hs.94211	cd1 (required for cell differentiation,	5.00	
	135103	NM_003428	Hs.9450	zinc finger protein 84 (HPF2)		11.00
	135145	AW014729	Hs.95262	nuclear factor related to kappa B bindin		4.01
	135184	U13222	Hs.96028	forkhead box D1		7.00
15	135242	AI583187	Hs.9700	cyclin E1	13.50	
	135286	AW023482	Hs.97849	ESTs	6.46	
	135289	AW372559	Hs.9788	hypothetical protein MGC10924 similar to		8.80
	135355	AK001652	Hs.99423	ATP-dependent RNA helicase	10.00	
20	135371	NM_006025	Hs.997	protease, serine, 22	8.00	
	135393	L11244	Hs.99886	complement component 4-binding protein,		14.60

TABLE 5B shows the accession numbers for those primekeys lacking unigenal D's for Table 5A. For each probe set we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probe set identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
117079	1621717_1	H92325 T97125
124305	242183_1	AW963221 AA344870 AA344871 H93331
101502	18202_6	M26958
109792	754958_1	R49625 F10674
126034	1598157_1	H60340 N91637
102768	44641_1	U82321 H66077
126345	1653833_1	N49713 NA9819 W03810
127066	1703458_1	R25066 R20144 R20145 Z43845
127099	244301_1	AA347668 AW956810 Z44271 F07065 F07064 R13506
119243	1774795_1	T12603 T12604
125875	1566433_1	H14480 N98295
112054	1538292_1	R43590 F10439
126979	171411_1	AA210954 AA211007
126992	880655_1	AI809521 H12174 Z42556
122318	292419_1	AA429743 AA442754
114699	135322_1	AA127386 R15644 AA127404
114793	150742_1	AA158245 AA158235
108305	111550_1	AA071391 AA069892 AA069891
108393	113411_1	AA075211 AA075245 AA075126 AA074946
100867	ligr_HT4586	U14622
123731	genbank_AA609839	AA609839
109700	genbank_F09609	F09609
120715	genbank_AA292700	AA292700
113702	genbank_T97307	T97307
115113	genbank_AA256460	AA256460
101045	entrez_J05614	J05614
108554	genbank_AA084948	AA084948
108573	genbank_AA086005	AA086005
119052	149538_1	R10889 R10888
126522	416020_1	W31912 AI167491
126605	439280_1	AA676910 AA778853 AA778865 W86800
103768	46922_1	W42667 AI580740 AI690440 AI561350 AW467906 AW151450 AI825927 AL041716 AI885600 AI742213 AW248624 AI955498 AA033947 AA845593 AI623711 N66583 C00064 AA193567 AW083868 AW163216 AA191595 AA522778 AI628008 AI915518 AA843508 AI926195 AA176265 AW167963 AA992115 W93647 AW103572 AI862994 AI342059 AA911719 AA176155 AA024712 AA069988 AA205591 AI591107 AI199673 AI811766 AI275832 AI422233 AI191852 AI096682 AI580124 AI683612 AA582453 AA927559 AA486415 T32414 AI084978 H44849 H44848 H20477 T91695 W47039 AA070055 AA024795 AA328855 AA379248 AA379330 AA385580 W25920 W03688 AA448359 AA093881 AW362477 AA089997 AI350265 W93479 N99688 AA932257 AW351469 H68590 AA663402 AA069771 AW087986 AI858420 AA600214 AI970774 AI857712 AI683081 AI885584 AW131150 AI567981 AW002714 AW189973 AW075495 AW168303 AA953714 AW516881 AI357375 AI566663 AW512676 AI570580 AI023690 AA448216 AI079853 AI422707 AA779516 AW026972 AW130082 AW162307 AW438646 AA709332 AW192394 AI167350 AI217879 AI129152 AA719509 AI350480 AA663418 AI003634 AW118546 AA180261 AA442833 AI268625 AA888881 AI038759 AA846723 AI248770 AA993694 AI280335 AI885107 AW518649 AA641563 AA995835 AA582521 AI276744 AA436478 AI017360 AI620763 AI859887 N73926 AI076327 AI741615 AI160617 AW172819 AI492005 AA677429 AA996334 AI693771 AI950039 AI245629 AI288515 AI866186 T93293 AI173262 AA599779 AI680092 AW439316 AI084555 AI272672 AI583507 AW473219 AA738132 AW473283 AI367492 AA995410 AI689624 AA206353 AI033095 AI040382 AA873630 AI221074 AI934840 AI418680 AA844306 R94503 AA773520 AA843169 AA219425 AA628658 AI811719 AW411275 AI590981 W37907 AI591178 AI684051 AA983238 AA669347 AA976239 AA704570 AI626339 AI684391 AI241580 AI003539 AW176687 AA009650 N34566 AI333493 AI186070 AA070827 AA411683 AI280884 AA872023 AA207255 AA021576 N71953 AI885888 AW076039 T15777 AI537673 AW248048 H09554 W93480 W47001 AW079114 AA063160 AA757453 R60788 AI859431 H20478 AA218882 AA757465 AA100995 AI64135 AI934209 AA070503 H47008 AA219646 W61039 W93907 AW385050 W37967 W78028 AA189007 AA479136 R93650 AA442312 T30287 AA847628 AA180262 AA009649 C03892 AW149464 AA310963 AA219693 AA069747 R29207 AA094784 AA293615 AA447848 AI984167 N90393 C05097 N56499 AW292351 AW149681 AW473258 AA629322 AI004409 AW105577 AI954937 AI811070 AA502422 AW514437 AA535460 AA916877 AW517122 AA974657 AA975649 AW517130 AW517129 F31737 W07688 AA193645 AA378994 AA489273 F32267 W39303 AA021181 N86810 AA406524 AA062553 AA436801 H08985 H15979 N40310

AA436789 AA232172 AW360778 W25862 R60282 AA436530 AA378894 AA187461 AI940535 AA604210 AA089514 AA360421 N88243 N84281  
 AA209340 N56174 N88374 AA191088 AW247691 AA249013 AA093111 AA972536 AW298594 AA375893 T12139 W28186 AW243849  
 A1288629 AA843996 W15260 A1188286 AW248079 R15836

119599 genbank\_W45552 W45552  
 112382 genbank\_R59904 R59904  
 105264 genbank\_AA227934 AA227934  
 100071 entrez\_A28102 A28102  
 123315 714071\_1 AA496369 AA496646

Table 6A shows 99 genes up-regulated nonsmokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59580 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

Key: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Unigene Title: Unigene gene title  
 R1: average of AI for samples from non-smokers with adenocarcinoma divided by the 90th percentile of AI for samples from smokers with adenocarcinoma  
 R2: average of AI for samples from non-smokers with squamous cell carcinoma divided by the 90th percentile of AI for samples from smokers with squamous cell carcinoma

Key	ExAccn	UnigenelD	Unigene Title	R1	R2
25	100971	BE379727	Hs.83213 fatty acid binding protein 4, adipocyte		3.64
	101174	L17330	Hs.280 pre-TANK cell associated protein	15.00	
	101296	Y12490	Hs.85092 thyroid hormone receptor interactor 11		2.46
	101304	AA001021	Hs.6685 thyroid hormone receptor interactor 8		12.00
	101806	AA586894	Hs.112408 S100 calcium-binding protein A7 (psorias		2.68
30	101972	S82472	gb:beta-pol-DNA polymerase beta (exon a		2.11
	102274	U30930	Hs.158540 UDP glycosyltransferase 8 (UDP-galactose	7.50	
	102394	NM_003816	Hs.2442 a disintegrin and metalloproteinase doma	7.50	
	102832	U92015	gb:Human clone 143789 defective mariner	13.50	
	103010	X52509	Hs.161640 tyrosine aminotransferase	9.50	
35	103439	X98266	gb:H.sapiens mRNA for ligase like protel		2.50
	103563	L02911	Hs.150402 activin A receptor, type I	9.00	
	103857	AJ076795	Hs.45033 lacrimal proline rich protein		3.94
	104239	AB002367	Hs.21355 doublecortin and CaM kinase-like 1	13.50	
	104590	AW373062	Hs.83623 nuclear receptor subfamily 1, group I, m		12.66
40	104907	AA055829	Hs.196701 ESTs, Weakly similar to ALU1_HUMAN ALU	16.50	
	106131	BE514788	Hs.296244 SNARE protein		2.17
	106672	H47233	Hs.30643 ESTs	7.00	
	106872	T56887	Hs.18282 KIAA1134 protein	11.50	
	106960	AA156238	Hs.32501 ESTs		2.38
45	106971	Z43846	Hs.194478 Homo sapiens mRNA; cDNA DKFZp434O1572 (f	9.50	
	107982	AA035375	Hs.57887 ESTs, Weakly similar to KIAA0758 protel		2.95
	108562	AA100796	Hs.292653 gb:zm26c06.s1 Stratagene pancreas (93720	16.50	
	108599	AB018549	Hs.69328 MD-2 protein	13.00	
	108663	BE219231	Hs.292653 ESTs, Weakly similar to T26845 hypotheti		2.40
50	109247	AA314907	Hs.85950 ESTs	7.00	
	109630	R44607	Hs.22672 ESTs		5.00
	110193	AJ004874	Hs.310764 Homo sapiens mRNA; cDNA DKFZp434M082 (fr	12.50	
	110234	H24458	Hs.32085 EST	16.50	
	110644	R94207	Hs.268989 ESTs, Highly similar to type II CALM/AF1	8.00	
55	110886	AW274992	Hs.72249 three-PDZ containing protein similar to	17.00	
	111057	T79639	Hs.14629 ESTs	16.50	
	111950	AF071594	Hs.110457 Wolf-Hirschhorn syndrome candidate 1	11.00	
	112291	R53972	Hs.26026 ESTs		3.00
	112956	Z43784	Hs.75893 ankyrin 3, node of Ranvier (ankyrin G)		2.79
60	113009	T23699	Hs.7246 ESTs		4.50
	113060	BE564162	Hs.250820 hypothetical protein FLJ14827	9.79	
	113073	N39342	Hs.103042 microtubule-associated protein 1B	32.50	
	113074	AK001335	Hs.31137 protein tyrosine phosphatase, receptor t		3.82
	113121	T48011	Hs.8764 EST		2.21
65	113125	AA568672	Hs.8929 hypothetical protein FLJ11362	19.50	
	113757	AA703095	Hs.18631 ESTs		2.65
	113848	W52854	Hs.27099 hypothetical protein FLJ23293 similar to	6.00	
	113884	AI333076	Hs.28529 chromosome 12 open reading frame 2		6.00
	113936	W17056	Hs.83623 nuclear receptor subfamily 1, group I, m		4.63
70	114875	AA235609	Hs.236443 Homo sapiens mRNA; cDNA DKFZp564N1063 (		7.00
	114987	AA251016	Hs.87808 EST		6.00
	115460	AW958439	Hs.38613 ESTs		2.27
	115722	W91892	Hs.59609 ESTs		9.00
	116261	AA481788	Hs.190150 ESTs	9.50	
75	116830	H61037	Hs.70404 ESTs, Weakly similar to ALU2_HUMAN ALU	8.50	
	116970	AB023179	Hs.9059 KIAA0962 protein	7.50	
	117178	H98675	Hs.269034 ESTs		2.68
	117757	AF088019	Hs.46732 EST	7.50	
	118283	AA287747	Hs.173012 ESTs, Weakly similar to A46010 X-linked	16.50	
80	118384	AF217525	Hs.49002 Down syndrome cell adhesion molecule		2.50
	118657	AJ822106	Hs.49902 ESTs		2.39
	120328	AA923278	Hs.290905 ESTs, Weakly similar to protease [H.sapi		3.50
	120404	AB023230	Hs.96427 KIAA1013 protein	7.00	
	120524	AA261852	Hs.192905 ESTs	6.00	
85	120688	AW207555	Hs.97093 Homo sapiens cDNA: FLJ23004 fis, clone L	17.82	

	121558	AA412497		gb:zt95g12.s1 Soares_testis_NHT Homo sap		2.95
	121676	H56037	Hs.108146	ESTs	10.00	
	121936	AI024600	Hs.98612	ESTs	15.00	
	121938	AA428659	Hs.98610	ESTs	14.00	
5	122177	AA435789	Hs.98833	EST	8.93	
	123442	AA299652	Hs.111496	Homo sapiens cDNA FLJ11643 fis, clone HE	13.04	
	123551	AA608837		gb:af03h12.s1 Soares_testis_NHT Homo sap	11.50	
	123756	AA609971	Hs.112795	EST	11.00	
10	123861	AA620840		gb:af89g01.s1 Soares_testis_NHT Homo sap		2.50
	124371	N24924	Hs.188601	ESTs	6.50	
	127477	BE326720	Hs.280651	ESTs		4.33
	127591	AI190540	Hs.131092	ESTs		3.02
	128262	AA456924	Hs.192228	ESTs	7.00	
	128426	AI265784	Hs.145197	ESTs		2.08
15	128925	R67419	Hs.21851	Homo sapiens cDNA FLJ12900 fis, clone NT		2.11
	128945	AI990506	Hs.8077	Homo sapiens mRNA; cDNA DKFZp547E184 (fr	10.00	
	129105	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	15.50	
	129235	AW977238	Hs.126084	KIAA1055 protein		4.25
	129506	AB020684	Hs.11217	KIAA0877 protein	6.50	
20	129595	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9		10.00
	130160	AA305688	Hs.267695	UDP-Gal4betaGlcNAc beta 1,3-galactosyltr	20.00	
	130340	D82326	Hs.239106	solute carrier family 3 (cystine, dibasi	11.50	
	131220	AB023194	Hs.300855	KIAA0977 protein	17.50	
	131430	AI879148	Hs.26770	fatty acid binding protein 7, brain	6.10	
25	132114	NM_006152	Hs.40202	lymphoid-restricted membrane protein		6.15
	132458	AA935315	Hs.48965	Homo sapiens cDNA: FLJ21693 fis, clone C		5.58
	132647	NM_006927	Hs.54432	siatyltransferase 4B (beta-galactosidase	7.50	
	132655	D49372	Hs.54460	small inducible cytokine subfamily A (Cy		2.53
	132682	AI077500	Hs.54900	serologically defined colon cancer antig		2.50
30	132747	AA345241	Hs.55950	ESTs, Weakly similar to KIAA1330 protein		2.83
	132812	R50333	Hs.92186	Leman coiled-coil protein		3.82
	133337	AF085983	Hs.293676	ESTs		5.00
	133876	AL134906	Hs.771	phosphorylase, glycogen; liver (Hers dis		3.00
	134119	AW157837	Hs.79226	fasciculation and elongation protein zet		2.06
35	134464	AA302983	Hs.239720	CCRA-NOT transcription complex, subunit		2.27
	134542	M14156	Hs.85112	insulin-like growth factor 1 (somatomedi		11.50
	135002	AA448542	Hs.251677	G antigen 7B	87.00	
	135305	AA203555	Hs.98288	Homo sapiens cDNA FLJ14903 fis, clone PL		6.50

TABLE 6B show the accession numbers for those primekeys lacking unigenelD's for Table 6A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey CAT number Accessions

108562	36375_1	AA100796	AF020589	AA074629	AA075946	AA100849	AA085347	AA126309	AA079311	AA079323	AA085274
103439	35330_1	X98266	N41124								
123551	genbank_AA608837		AA608837								
123861	genbank_AA620840		AA620840								
102832	entrez_U92015		U92015								
101972	entrez_S82472		S82472								
121558	genbank_AA412497		AA412497								



Table 7A shows 98 genes down-regulated in non-smokers with lung cancer relative to smokers with lung cancer. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5	Pkey:	Unique Eos probeset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigenetID:	Unigene number			
	Unigene Title:	Unigene gene title			
10	R1:	90th percentile of AI for samples from smokers with adenocarcinoma divided by the average of AI for samples from non-smokers with adenocarcinoma.			
	R2:	90th percentile of AI for samples from smokers with squamous cell carcinoma divided by the average of AI for samples from non-smokers with squamous cell carcinoma.			
	Pkey	ExAccn	UnigenetID	Unigene Title	R1 R2
15	100187	D17793	Hs.78183	aldo-keto reductase family 1, member C3	164.10
	100380	D82343	Hs.18551	neuroblastoma (nerve tissue) protein	77.40
	100576	X00356	Hs.37058	calitonin/calcitonin-related polypeptid	102.40
	100971	BE379727	Hs.83213	fatty acid binding protein 4, adipocyte	463.80
	101046	K01160		(NONE)	672.00
20	101066	AW970254	Hs.889	Charol-Leyden crystal protein	66.00
	101175	U82671	Hs.36980	melanoma antigen, family A, 2	77.20
	101497	W05150	Hs.37034	homeo box A5	62.80
	101663	NM_003528	Hs.2178	H2B histone family, member Q	78.00
25	101677	NM_000715	Hs.1012	complement component 4-binding protein,	186.20
	101745	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	80.08
	101941	S75583		gb:HERVK10/HUMMTV reverse transcriptase	99.20
	102125	NM_006456	Hs.288215	sialyltransferase	103.10
	102242	U27185	Hs.82547	retinoic acid receptor responder (tazaro	67.00
30	102340	U37055	Hs.278657	macrophage stimulating 1 (hepatocyte gro	71.60
	102369	U39840	Hs.299867	hepatocyte nuclear factor 3, alpha	69.70
	102457	NM_001394	Hs.2359	dual specificity phosphatase 4	153.00
	102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	65.70
	102796	AL079846	Hs.107019	symplesin; Huntingtin interacting protei	58.80
35	102829	NM_006183	Hs.80962	neurotensin	268.80
	103207	X72790		gb:Human endogenous retrovirus mRNA for	70.00
	103242	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	212.10
	103260	X78416	Hs.3155	casein, alpha	130.70
	103351	X89211		gb:H.sapiens DNA for endogenous retrovir	64.60
40	104212	AB002298	Hs.173035	KIAA0300 protein	66.80
	104252	AF002246	Hs.210863	cell adhesion molecule with homology to	63.80
	104258	AF007216	Hs.5462	solute carrier family 4, sodium bicarbon	94.40
	105024	AA126311	Hs.9879	ESTs	68.20
	106260	AI097144	Hs.5250	ESTs, Weakly similar to ALU1_HUMAN ALU S	74.60
45	106440	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	71.10
	106566	BE298210		gb:501118016F1 NIH_MGC_17 Homo sapiens c	73.20
	106605	AW772298	Hs.21103	Homo sapiens mRNA; cDNA DKFZp564B076 (fr	83.80
	106614	AA648459	Hs.335951	hypothetical protein AF301222	62.30
	106654	AW075485	Hs.285049	phosphoserine aminotransferase	202.40
50	106999	H93281	Hs.10710	hypothetical protein FLJ20417	89.60
	108700	AA121518	Hs.193540	ESTs, Moderately similar to 2109260A B c	66.40
	108810	AW295647	Hs.71331	hypothetical protein MGC6350	95.60
	108857	AK001468	Hs.62180	anillin (Drosophila Scraps homolog), act	63.40
	109597	AA989362	Hs.283780	ESTs	85.00
55	109691	T65568	Hs.12860	ESTs	58.70
	109704	AI743880	Hs.12876	ESTs	60.60
	110942	R63503	Hs.28419	ESTs	76.40
	111722	R23924	Hs.23596	EST	74.60
	112891	T03927	Hs.283147	ESTs, Moderately similar to A46010 X-II	64.80
60	112992	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp781J1324 (f	76.70
	113073	N39342	Hs.103042	microtubule-associated protein 1B	120.20
	114251	H15261	Hs.21948	ESTs	127.20
	115230	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fis, clone L	174.00
	115291	BE545072	Hs.122579	hypothetical protein FLJ10461	91.00
65	115815	AW905328	Hs.180842	ribosomal protein L13	66.40
	115909	AW872527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	226.60
	115965	AA001732	Hs.173233	hypothetical protein FLJ10970	82.80
	116107	AL133916	Hs.172572	hypothetical protein FLJ20093	361.60
	116552	D20508	Hs.164649	hypothetical protein DKFZp434H247	69.00
	116571	D45652		gb:HUMGSO2848 Human adult lung 3' direct	64.20
70	118466	N65741		gb:yz33g08.s1 Morton Fetal Cochlea Homo	63.50
	120484	AA253170	Hs.96473	EST	81.60
	120983	AA398209	Hs.97587	EST	81.10
	121034	AL389951	Hs.271623	nucleoporin 50kD	66.20
75	121423	AW873352	Hs.290585	ESTs	64.40
	122553	AA451884	Hs.190121	ESTs	60.40
	122946	AI718702	Hs.308026	major histocompatibility complex, class	189.60
	123130	AA487200		gb:ab19f02.s1 Stratagene lung (937210) H	80.20
	124472	N52517	Hs.102670	EST	71.00
80	124526	N62096	Hs.293185	ESTs, Weakly similar to JCT328 amino aci	104.80
	125489	H49193	Hs.124984	ESTs, Moderately similar to ALU7_HUMAN A	72.00
	125731	R61771	Hs.26912	ESTs	69.90
	125747	NM_002884	Hs.865	RAP1A, member of RAS oncogene family	69.00
	126020	H79863	Hs.114243	ESTs	62.40
	126547	U47732	Hs.84072	transmembrane 4 superfamily member 3	62.80
85	126966	R38438	Hs.182575	solute carrier family 15 (H+peptide tra	60.10

	127472	AA761378	Hs.192013	ESTs	70.20	
	127610	AA960867	Hs.160271	ESTs, Highly similar to unnamed protein	64.00	
	127742	AW293496	Hs.180138	ESTs	85.20	
	127987	AI022103	Hs.124511	ESTs	96.60	
5	128233	AW889132	Hs.11916	ribokinase		78.90
	128420	AA650274	Hs.41296	fibronectin leucine rich transmembrane p		108.90
	128766	AW160432	Hs.296460	craniofacial development protein 1	66.80	
	129014	AW935187	Hs.170162	KIAA1357 protein		58.53
10	129215	AB040930	Hs.126085	KIAA1497 protein	64.20	
	130090	H97878	Hs.132390	zinc finger protein 36 (KIX 18)	63.80	
	130385	AW067800	Hs.155223	slanniocalcin 2		139.60
	130732	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)		64.60
	131025	AB040900	Hs.6189	KIAA1467 protein	64.40	
15	131241	BE501914	Hs.24654	Homo sapiens cDNA FLJ11640 fis, clone HE	76.20	
	131775	AB014548	Hs.31921	KIAA0548 protein	97.80	
	132240	AB018324	Hs.42676	KIAA0781 protein		71.00
	132856	NM_001448	Hs.58367	glypican 4		88.40
	132977	AA093322	Hs.301404	RNA binding motif protein 3	133.20	
20	133749	L20852	Hs.10018	solute carrier family 20 (phosphate tran		59.30
	133818	AI110684	Hs.7645	fibrinogen, B beta polypeptide	341.00	
	134264	AF149297	Hs.8087	NAG-5 protein		64.30
	134265	M83772	Hs.80876	flavin containing monooxygenase 3		232.53
	134346	X84002	Hs.82037	TATA box binding protein (TBP)-associate	66.00	
25	134395	AA456539	Hs.8282	lysosomal-associated membrane protein 2		75.80
	135047	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su		108.30
	135056	N75765	Hs.93765	lipoma HMGIC fusion partner	71.40	
	135309	AI564123	Hs.42500	ADP-ribosylation factor-like 5	70.40	

30 TABLE 7B shows the accession numbers for those primekeys lacking unigenelD's for Table 7A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

35 Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
103207	30635_-4	X72790
106566	120358_1	BE288210 AI672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AI287650 AI885299 AI985381 AW592624 AW340136 AI266556 AA456390 AI310815 AA484951
116571	genbank_D45652	D45652
118466	genbank_N66741	N66741
101046	entrez_K01160 K01160	
101941	entrez_S77583 S77583	
103351	entrez_X89211 X89211	
123130	genbank_AA487200	AA487200

Table 8A shows 1720 genes either up or down-regulated in lung tumors or chronically diseased lung relative to a broad collection of over 40 distinct normal body tissues. Chronically diseased lung samples represent chronic non-malignant lung diseases such as fibrosis, emphysema, and bronchitis. These genes were selected from 39494 probesets on the Eos/Affymetrix Hu02 Genechip array. Gene expression data for each probeset obtained from this analysis was expressed as average intensity (AI), a normalized value reflecting the relative level of mRNA expression.

5  
Pkey: Unique Eos probeset identifier number  
ExAccn: Exemplar Accession number, Genbank accession number  
UnigeneID: Unigene number  
Unigene Title: Unigene gene title  
10 R1: 70th percentile of AI for lung tumors divided by 90th percentile of AI for normal lung  
R2: 70th percentile of AI for chronically diseased lung divided by 90th percentile of AI for normal lung

	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
15	300097	AI916973	Hs.213603	ESTs	5.46	4.69
	300117	AW189787	Hs.147474	ESTs	0.58	0.56
	300197	AI686661	Hs.218286	ESTs	4.26	5.44
	300201	AI308300		gb:ta90c06x1 NCI_CGAP_Bm20 Homo sapien	0.62	0.83
	300225	AI989963	Hs.197505	ESTs	1.68	1.75
20	300247	AW274682	Hs.161394	ESTs	1.08	2.28
	300258	AI469095	Hs.298241	Transmembrane protease, serine 3	0.86	1.00
	300337	AI707881	Hs.202090	ESTs	5.80	9.09
	300362	Z42308		gb:HSC0FB121 normalized infant brain cDN	4.18	12.78
	300374	AI859947	Hs.314158	ESTs	2.99	4.38
25	300387	AW270150	Hs.254516	ESTs	1.50	2.53
	300440	AI421541	Hs.146184	ESTs	3.98	5.25
	300441	R10367	Hs.307921	EST, Weakly similar to Z232_HUMAN ZINC F	3.18	6.80
	300449	AI362967	Hs.132221	hypothetical protein FLJ12401	0.43	0.62
	300469	AW135830	Hs.233955	hypothetical protein FLJ20401	0.16	0.83
30	300552	X85711	Hs.21838	hypothetical protein FLJ11191	4.10	9.75
	300627	W27363		gb:ab37d01.r1 Stratagene HeLa cell s3 93	4.60	12.60
	300630	AW118822	Hs.128757	ESTs	2.91	5.86
	300716	AI216113	Hs.126280	hypothetical protein FLJ23393	1.00	0.92
	300738	AI623332	Hs.130541	KIAA1542 protein	1.82	1.71
35	300777	AA235361	Hs.96840	KIAA1527 protein	4.48	8.22
	300790	AI492471	Hs.188270	ESTs	1.29	1.18
	300832	AI688147	Hs.220615	ESTs, Weakly similar to T03829 transcrip	5.51	8.56
	300836	Z44942	Hs.22958	calcium channel alpha2-delta3 subunit	4.80	6.34
	300838	AI582897	Hs.192570	hypothetical protein FLJ22028	1.70	2.81
40	300878	AW449802	Hs.285901	Homo sapiens cDNA FLJ20428 fis, clone KA	4.56	7.91
	300897	AI890356	Hs.127804	ESTs, Weakly similar to T17233 hypotheti	2.23	1.58
	300926	AA504860		gb:ab03a10.s1 Stratagene fetal retina 93	2.13	3.50
	300960	AI041019	Hs.152454	ESTs	2.74	4.46
	300961	AW204069	Hs.312716	ESTs, Weakly similar to unnamed protein	1.00	1.00
45	300962	AA593373	Hs.293744	ESTs	1.46	1.51
	300967	AA565209	Hs.269439	ESTs	0.39	1.30
	300987	AW450840	Hs.148590	ESTs, Weakly similar to AF208846 1 BM-00	1.49	1.08
	300988	AI927208	Hs.208952	ESTs	0.16	0.37
	301050	AW136973	Hs.268516	ESTs, Weakly similar to S69890 mitogen i	3.23	1.94
50	301098	AA677570	Hs.185918	ESTs	6.76	14.28
	301157	AA729905	Hs.231916	ESTs	3.16	8.85
	301162	AI142118	Hs.129004	ESTs	1.68	7.18
	301170	AA737594	Hs.247606	ESTs	4.40	6.42
	301192	AI808751	Hs.121188	ESTs	6.38	11.59
55	301193	AA758115	Hs.128350	ESTs, Weakly similar to JC5423 2-hydroxy	4.35	7.78
	301267	AW297762	Hs.255690	ESTs	1.56	1.61
	301281	AA843986	Hs.190586	ESTs	2.19	1.78
	301341	AI819198	Hs.208229	ESTs	0.76	0.76
	301382	AA912839	Hs.163369	ESTs	1.00	1.81
60	301407	AW450466	Hs.126830	ESTs	1.48	1.51
	301452	AA975688	Hs.159955	ESTs	0.51	1.46
	301483	AW272467	Hs.254655	Unlabeled	2.40	5.02
	301494	AI678034	Hs.131099	ESTs	2.79	3.41
65	301521	AI733621	Hs.133011	zinc finger protein 117 (HPF9)	0.67	0.67
	301531	AI077462	Hs.134084	ESTs	2.52	3.76
	301580	AI878959	Hs.73737	splicing factor, arginina/serine-rich 1	7.41	11.92
	301676	Z43570	Hs.27453	ESTs, Moderately similar to G01251 Rar p	8.31	10.70
	301690	F05865	Hs.108323	ubiquitin-conjugating enzyme E2E 2 (homo	2.70	4.22
	301718	F07744	Hs.7987	DKFZP434F162 protein	4.20	8.78
70	301799	AA384252	Hs.286132	D15F37 (pseudogene)	5.93	7.04
	301804	AA581004	Hs.62180	anillin (Drosophila Scraps homolog), act	1.70	0.76
	301822	X17033	Hs.271986	Integrin, alpha 2 (CD49B, alpha 2 subuni	1.58	1.36
	301846	R20002	Hs.6823	hypothetical protein FLJ10430	1.00	1.00
	301868	T71508	Hs.13861	ESTs, Weakly similar to pH sensitive max	2.88	6.49
75	301882	T78054		gb:yc57g09.r1 Soares infant brain 1N1B H	2.28	3.80
	301905	AI991127	Hs.117202	ESTs	1.00	1.00
	301948	AA344647	Hs.116724	aldo-keto reductase family 1, member B11	5.28	2.28
	301960	AW070252	Hs.27973	KIAA0874 protein	5.38	6.48
	302011	T91418	Hs.125166	transcriptional adaptor 2 (ADA2, yeast,	3.03	3.42
80	302016	N40834	Hs.23495	hypothetical protein FLJ11252	1.00	1.25
	302041	NM_001501	Hs.129715	gonadotropin-releasing hormone 2	0.71	0.99
	302072	AJ238381	Hs.132576	paired box gene 9	1.60	1.71
	302094	AI286176	Hs.6786	ESTs	0.52	1.20
	302095	AW044300	Hs.137506	Homo sapiens BAC clone RP11-120J2 from 7	2.75	4.93
85	302148	AW269618	Hs.23244	ESTs	3.04	3.87

	302155	AI088485	Hs.144759	ESTs	0.45	1.15
	302201	AJ006276	Hs.159003	transient receptor potential channel 6	0.33	0.84
	302202	AF097159	Hs.159140	UDP-Gal:betaGlcNAc beta 1,4- galactosylt	0.52	0.94
5	302206	AI937193	Hs.41143	phosphoinositide-specific phospholipase	2.76	3.65
	302209	AF047445	Hs.159297	killer cell lectin-like receptor subfami	1.00	1.00
	302235	AL049987	Hs.186361	Homo sapiens mRNA; cDNA DKFZp564F112 (fr	1.68	1.50
	302290	AL117607	Hs.175563	Homo sapiens mRNA; cDNA DKFZp564N0763 (f	1.00	2.11
	302328	AA354849	Hs.23240	Homo sapiens cDNA FLJ13496 fis, clone PL	9.38	13.08
10	302346	AL039101	Hs.194825	dynein, cytoplasmic, light intermediate	3.27	7.24
	302360	AJ010901	Hs.198267	mucin 4, tracheobronchial	2.54	1.88
	302384	Y08982	Hs.202676	synaptonemal complex protein 2	1.00	0.91
	302406	U86751	Hs.211956	CD3-epsilon-associated protein; antisens	2.63	2.67
	302409	AF155156	Hs.218028	adaptor-related protein complex 4, epsil	5.82	9.34
15	302423	AB028977	Hs.225974	KIAA1054 protein	3.66	3.18
	302432	AL080058	Hs.272534	Homo sapiens mRNA; cDNA DKFZp564J062 (fr	2.44	6.77
	302435	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolo	0.44	0.84
	302437	AB024730	Hs.227473	UDP-N-acetylglucosaminase-1,3-D-mannosid	4.18	5.64
	302455	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	1.85	0.92
20	302472	AA317451	Hs.24335	SW/SNF related, matrix associated, acti	2.04	2.13
	302476	AF182234	Hs.241578	U6 snRNA-associated Sm-like protein LSm8	1.44	1.89
	302489	T80660	Hs.230424	Homo sapiens cDNA FLJ13540 fis, clone PL	0.51	1.10
	302490	AA885502	Hs.187032	ESTs	2.64	4.87
	302562	AJ005585	Hs.48956	gap junction protein, beta 6 (connexin 3	5.34	2.68
25	302566	AA085996	Hs.248572	hypothetical protein FLJ22965	1.00	1.21
	302630	AB029488	Hs.272100	SMS3 protein	0.52	1.24
	302634	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	1.00	1.00
	302638	AA463798	Hs.102696	MCT-1 protein	1.58	1.02
	302647	X57723	Hs.198273	NADH dehydrogenase (ubiquinone) 1 beta s	2.72	6.85
30	302655	AJ227892	Hs.146274	ESTs	1.00	4.32
	302656	AW293005	Hs.70704	Homo sapiens, clone IMAGE:2823731, mRNA,	2.97	0.93
	302668	AA580691	Hs.180789	S164 protein	0.80	0.95
	302679	H65022		gb:yu66g11.1 r1 Weizmann Olfactory Epithel	1.68	5.04
	302680	AW192334	Hs.38218	ESTs	2.70	7.98
35	302697	AJ001408		gb:Homo sapiens mRNA for immunoglobulin	4.25	8.13
	302705	U09060		gb:Human immunoglobulin heavy chain, V-r	3.91	8.68
	302711	L08442		gb:Human autonomously replicating sequen	2.20	2.73
	302719	W69724	Hs.268959	hypothetical protein FLJ20920	0.54	1.02
	302742	L12069		gb:Homo sapiens (clone WR4.10VH) anti-th	4.28	11.57
40	302755	AW384815	Hs.149208	KIAA1655 protein	1.57	2.38
	302771	H98476	Hs.42522	ESTs	2.94	4.68
	302769	AJ245067		gb:Homo sapiens mRNA for immunoglobulin	3.49	6.31
	302795	AJ245313	Hs.272838	hypothetical protein FLJ10494	0.80	2.74
	302802	Y08250		gb:H.sapiens mRNA for variable region of	1.13	0.77
45	302803	AA442824	Hs.293961	ESTs, Moderately similar to putative DNA	3.14	10.68
	302812	N31301	Hs.152664	hypothetical protein FLJ20051	3.04	8.24
	302847	X98940		gb:H.sapiens rearranged Ig heavy chain (	1.80	1.92
	302865	AL137763	Hs.132127	hypothetical protein LOC57822	1.00	1.00
	302943	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheri	0.53	0.67
50	302977	AW263124	Hs.315111	hypothetical protein FLJ12894	2.45	2.62
	303008	AF078950	Hs.24139	Homo sapiens cDNA: FLJ23137 fis, clone L	4.88	8.61
	303011	AF090405		gb:Homo sapiens clone 2A1 scFV antibody	1.41	1.86
	303013	F07898	Hs.288968	RAB22A, member RAS oncogene family	1.51	1.19
55	303061	AF151882	Hs.27693	peptidylprolyl isomerase (cyclophilin)-I	0.72	0.76
	303077	AF163305		gb:H.sapiens T-cell receptor mRNA	1.17	3.90
	303080	AA443259	Hs.146286	kinesin family member 13A	4.08	6.46
	303091	AF192913	Hs.130683	zinc finger protein 180 (HHZ168)	2.50	4.37
	303094	AF195513	Hs.278953	Pur-gamma	5.38	8.38
	303095	AF202051	Hs.134079	NM23-H8	3.26	4.08
60	303131	AW081081	Hs.103180	OC2 protein	2.02	1.83
	303195	AA082211	Hs.233936	myosin, light polypeptide, regulatory, n	1.32	3.95
	303196	AA082298	Hs.59710	ESTs	0.77	0.53
	303216	AA581439	Hs.152328	ESTs	0.24	0.63
	303222	AA333538	Hs.204501	hypothetical protein FLJ10534	3.56	6.22
65	303234	AA132255	Hs.143951	ESTs	2.28	3.17
	303251	AW340037	Hs.115897	protocadherin 12	0.38	1.02
	303295	AA205625	Hs.208067	ESTs	2.30	1.00
	303297	T80072	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.86	4.48
70	303316	AF033122	Hs.14125	p53 regulated PA26 nuclear protein	0.10	0.80
	303467	AA398801	Hs.323397	ESTs	4.54	9.65
	303508	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	0.09	0.04
	303552	AA359799	Hs.224662	ESTs, Weakly similar to unnamed protein	1.00	1.72
	303598	AA382814		gb:EST96097 Testis I Homo sapiens cDNA 5	4.96	9.14
	303637	AF056083	Hs.24879	phosphatidic acid phosphatase type 2C	2.05	2.02
75	303655	AA504702	Hs.258802	ATPase, (Na+)/K+ transporting, beta 4 po	1.00	1.24
	303756	AI738488	Hs.115838	ESTs	1.08	1.43
	303856	AA968589	Hs.180532	glucose phosphate isomerase	1.76	1.31
	303893	N88597	Hs.113503	karyopherin (importin) beta 3	2.30	2.57
	303907	AW467774	Hs.171880	polymerase (RNA) II (DNA directed) polyp	3.10	5.79
80	303946	AW474196	Hs.306637	Homo sapiens cDNA FLJ12363 fis, clone MA	5.06	11.86
	303978	AW513315		gb:bxo43c12.x1 NCL_CGAP_Ut1 Homo sapiens	5.14	7.31
	303981	AW513804	Hs.278834	ESTs, Weakly similar to ALU1_HUMAN ALU S	2.83	4.06
	303990	AW515465		gb:bxo71a11.x1 NCL_CGAP_Kd8 Homo sapiens	1.15	2.35
	303998	AW516449		gb:bxo58f05.x1 NCL_CGAP_Ut2 Homo sapiens	2.20	9.35
85	303999	AW516611		gb:bxo70b11.x1 NCL_CGAP_Ov39 Homo sapiens	4.85	6.28
	304006	AW517947		gb:bxo66h02.x1 NCL_CGAP_Ut2 Homo sapiens	3.21	4.07

	304008	AW518198	Hs.3297	ribosomal protein S27a	6.50	11.08
	304009	AW518206	Hs.181165	eukaryotic translation elongation factor	1.88	3.27
	304024	T03036		gb:FB2187 Fetal brain, Stratagene Homo s	2.15	3.55
5	304026	T03160		gb:FB26F2 Fetal brain, Stratagene Homo s	5.88	11.80
	304028	T03266		gb:FB7C1 Fetal brain, Stratagene Homo sa	5.59	13.46
	304036	T16855	Hs.244621	ribosomal protein S14	6.55	14.43
	304046	T54803		gb:yb42d08.s1 Stratagene fetal spleen (9	6.18	12.19
	304061	T61521		gb:yb73g01.s1 Stratagene ovary (937217)	2.64	8.23
10	304063	T62536		gb:yc04c12.s1 Stratagene lung (937210) H	0.53	1.61
	304097	R25376	Hs.177592	ribosomal protein, large, P1	6.49	11.67
	304114	R78946		gb:yl87g02.s1 Soares placenta Nb2HP Homo	2.90	4.18
	304122	H28966		gb:ym31a05.s1 Soares infant brain 1N1B H	1.00	2.76
	304155	H68696		gb:yr78b05.s1 Soares fetal liver spleen	0.79	1.18
15	304203	N56929		gb:yy82d08.s1 Soares_mulliple_sclerosis_	4.28	11.34
	304234	W81608		gb:zd88h06.s1 Soares_fetal_heart_NbHH19W	6.47	11.03
	304267	AA064862	Hs.73742	ribosomal protein, large, P0	1.34	1.16
	304270	AA069711	Hs.297753	vimentin	3.40	5.40
	304287	AA079286	Hs.78466	proteasome (prosome, macropain) 26S sub	2.93	4.42
20	304348	AA179868		gb:zp38g12.s1 Stratagene muscle 937209 H	3.98	10.96
	304415	AA250747	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	3.32	5.99
	304430	AA347682		gb:EST54044 Fetal heart II Homo sapiens	1.00	1.00
	304456	AA411240		gb:zv26g05.s1 Soares_NhHMPu_S1 Homo sapi	1.42	3.33
	304521	AA646716		gb:zx82c11.s1 Soares ovary tumor NbHOT H	2.18	1.15
25	304526	AA676427		gb:zx02c05.s1 Soares_total_fetus_Nb2HFB_	5.38	14.11
	304542	AA482602	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.16	8.23
	304545	AA486074	Hs.297681	serine (or cysteine) proteinase inhibitor	0.55	1.20
	304607	AA513322		gb:nh85e08.s1 NCI_CGAP_Br1.1 Homo sapien	1.95	2.10
	304640	AA524440	Hs.111334	ferritin, light polypeptide	2.10	2.83
30	304650	AA527489	Hs.3463	ribosomal protein S23	3.33	12.62
	304735	AA576453		gb:nm75h11.s1 NCI_CGAP_Co9 Homo sapiens	1.33	0.88
	304760	AA580401		gb:nm13g09.s1 NCI_CGAP_Co12 Homo sapiens	3.68	8.14
	304849	AA588157	Hs.13801	KIAA1685 protein	2.77	3.70
	304917	AA602685	Hs.284136	PRO2047 protein	7.16	11.01
35	304921	AA603092	Hs.297753	vimentin	2.47	4.24
	304966	AA613893	Hs.282435	ESTs	6.78	11.66
	304987	AA618044	Hs.300697	immunoglobulin heavy constant gamma 3 (G	0.90	1.23
	305016	AA626876		gb:zu69h05.s1 Soares_testis_NHT Homo sap	6.46	10.17
	305034	AA630128		gb:zb99c04.s1 Stratagene lung (937210) H	1.00	1.00
40	305072	AA641012		gb:mr72a12.s1 NCI_CGAP_Pr24 Homo sapiens	5.68	11.59
	305111	AA644187	Hs.303405	ESTs	1.48	1.37
	305148	AA654070		gb:nl01g08.s1 NCI_CGAP_Lym3 Homo sapiens	1.76	4.61
	305159	AA659166	Hs.275668	EST, Weakly similar to EF1D_HUMAN ELONG	1.00	2.15
	305190	AA659555		gb:ag57d12.s1 Gessler Wilms tumor Homo s	5.31	8.14
45	305232	AA670052	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	0.78	1.18
	305235	AA670480		gb:ag37e01.s1 Jia bone marrow stroma Hom	3.11	8.66
	305245	AA676695	Hs.81328	nuclear factor of kappa light polypeptid	4.38	7.53
	305312	AA700201		gb:z44f07.s1 Soares_fetal_liver_spleen_	2.13	2.66
	305322	AA701597	Hs.163019	EST	1.20	1.40
50	305394	AA720942	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.16	0.68
	305413	AA724659		gb:al10f08.s1 Soares_parathyroid_tumor_N	5.86	9.87
	305447	AA737856		gb:mx10c08.s1 NCI_CGAP_GC3 Homo sapiens	2.21	2.86
	305476	AA745664	Hs.287445	hypothetical protein FLJ11728	3.36	6.54
	305483	AA748030	Hs.303512	EST	1.00	2.02
55	305528	AA769156		gb:znz12e05.s1 NCI_CGAP_GCB1 Homo sapiens	6.44	9.10
	305612	AA782347	Hs.272572	hemoglobin, alpha 2	0.19	0.79
	305614	AA782866		gb:aj09h02.s1 Soares_parathyroid_tumor_N	1.00	1.00
	305616	AA782884	Hs.275865	ribosomal protein S18	7.57	10.20
	305637	AA806124		gb:oe29a12.s1 NCI_CGAP_Pr25 Homo sapiens	4.78	12.42
60	305639	AA806138		gb:oe29c12.s1 NCI_CGAP_Pr25 Homo sapiens	0.89	0.70
	305650	AA807709		gb:mw31e04.s1 NCI_CGAP_GC80 Homo sapiens4.49	4.91	8.71
	305690	AA813477		gb:al67a05.s1 Soares_testis_NHT Homo sap	4.91	9.40
	305726	AA828156	Hs.73742	ribosomal protein, large, P0	0.19	0.81
	305728	AA828209		gb:of34a02.s1 NCI_CGAP_Kid6 Homo sapiens	5.12	9.29
65	305759	AA835353		gb:ak72b06.s1 Barstead spleen HPLRB2 Hom	1.66	4.11
	305792	AA845256		gb:ak84a08.s1 Barstead spleen HPLRB2 Hom	2.34	4.25
	305864	AA864374	Hs.73742	ribosomal protein, large, P0	0.30	1.40
	305901	AA872958		gb:oh63h08.s1 NCI_CGAP_Kid5 Homo sapiens	2.10	5.21
	305910	AA875981		gb:mx21h02.s1 NCI_CGAP_GC3 Homo sapiens	0.32	1.01
70	306015	AA897116		gb:am08b07.s1 Soares_NFL_T_GBC_S1 Homo s1.56	1.12	1.12
	306017	AA897221	Hs.109058	ribosomal protein S6 kinase, 80kD, polyp	5.21	7.90
	306020	AA897630	Hs.130027	EST	1.96	6.59
	306063	AA906316		gb:ok03g03.s1 Soares_NFL_T_GBC_S1 Homo s	7.38	20.69
	306065	AA906725		gb:ok78g02.s1 NCI_CGAP_GC4 Homo sapiens	7.19	13.48
75	306104	AA910956		gb:ok85h11.s1 NCI_CGAP_Kid3 Homo sapiens	6.50	9.13
	306109	AA911651		gb:og21a07.s1 NCI_CGAP_PNS1 Homo sapiens	4.21	5.25
	306148	AA917409	Hs.288036	tRNA isopentenylpyrophosphate transferase	2.20	2.70
	306242	AA932805		gb:oo60g04.s1 NCI_CGAP_Lu5 Homo sapiens	2.84	5.35
	306288	AA936900		gb:ol53h05.s1 NCI_CGAP_HN3 Homo sapiens	1.60	1.12
80	306325	AA953072	Hs.210546	interleukin 21 receptor	1.65	2.26
	306353	AA961382	Hs.275965	ribosomal protein S18	3.78	6.32
	306375	AA968650	Hs.276018	EST, Moderately similar to JC4652 ribos	4.30	5.74
	306396	AA970223		gb:op09d05.s1 NCI_CGAP_Kid6 Homo sapiens	0.95	2.45
	306428	AA975110	Hs.191228	hypothetical protein FLJ20284	3.19	4.10
85	306442	AA976899		gb:oo35e09.s1 NCI_CGAP_GC4 Homo sapiens	4.67	7.44
	306446	AA977348		gb:ooq72a12.s1 NCI_CGAP_Kid6 Homo sapiens	3.92	6.27

5	306458	AA978186		gbxp33c06.s1 Soares_NFL_T_GBC_S1 Homo s	3.35	5.77
	306467	AA983508	Hs.163593	ribosomal protein L18a	3.72	5.37
	306510	AA988546		gbx0r84d07.s1 NCL_CGAP_Lu5 Homo sapiens	1.00	1.00
	306555	AA994304	Hs.276083	EST, Weakly similar to RL23_HUMAN 60S R	6.61	10.91
	306557	AA994530		gbxou57e08.s1 NCL_CGAP_Br2 Homo sapiens	16.20	31.83
10	306572	AA995686		gbxos25c12.s1 NCL_CGAP_Kid5 Homo sapiens	2.51	6.52
	306582	AA996248		gbxos18c10.s1 NCL_CGAP_Kid5 Homo sapiens	1.42	3.13
	306598	A1000320	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.91	8.68
	306605	A1000497	Hs.119500	ribosomal protein, large P2	1.96	8.60
	306656	A1004024		gbxou11b07.x1 Soares_NFL_T_GBC_S1 Homo s	0.11	0.45
15	306676	A1005603	Hs.284136	PRO2047 protein	9.56	17.28
	306686	A1015615		gbxov29f10.x1 Soares_testis_NHT Homo sap	1.86	3.60
	306702	A1022565	Hs.307670	EST	1.47	1.19
	306728	A1027359	Hs.272572	hemoglobin, alpha 2	1.28	2.83
	306751	A1032589		gbxow70h12.s1 Soares_fetal_liver_spleen_	3.91	5.21
20	306767	A1038963	Hs.249118	ESTs	3.33	6.06
	306892	A1092465		gbxqa75h12.x1 Soares_fetal_heart_NbHH19W	3.77	7.46
	306897	A1093967		gbxqa33c06.s1 Soares_NhHMPu_S1 Homo sapi	2.12	2.85
	306956	A1125111		gbxam66f03.s1 Barstead spleen HPLRB2 Hom	6.10	10.52
	306958	A1125152		gbxam55e09.x1 Johnston frontal cortex Ho	1.72	1.56
25	307035	A1142774	Hs.119122	ribosomal protein L13a	2.00	4.70
	307041	A1144243		gbxqb85b12.x1 Soares_fetal_heart_NbHH19W	9.12	12.56
	307091	A1167439		gbxox70h06.s1 Soares_NhHMPu_S1 Homo sapi	4.88	8.52
	307181	A1189251		gbxqc99g06.x1 Soares_pregnant_uterus_NbH	3.55	6.44
	307297	A1205799	Hs.111334	fertilin, light polypeptide	2.46	4.65
30	307317	A1208303	Hs.147333	EST	5.64	10.13
	307327	A1214142	Hs.246381	CD68 antigen	3.18	5.15
	307382	A1223158	Hs.147885	ESTs	2.02	3.73
	307410	A1241715	Hs.77039	ribosomal protein S3A	0.72	0.48
	307415	A1242118		gbxqh92b02.x1 Soares_NFL_T_GBC_S1 Homo s	2.38	3.51
35	307423	A1243206	Hs.179573	collagen, type I, alpha 2	2.60	5.44
	307426	A1243364		gbxqh30g11.x1 Soares_NFL_T_GBC_S1 Homo s	3.18	7.67
	307517	A1275055		gbxql72d03.x1 Soares_NhHMPu_S1 Homo sapi	1.00	1.00
	307551	A1281556		gbxqu52f11.x1 NCL_CGAP_Lym6 Homo sapiens	3.40	11.20
	307561	A1282207		gbxqp65a12.x1 Soares_fetal_lung_NbHL19W	4.74	15.51
40	307608	A1290295		gbxqm01f02.x1 Soares_NhHMPu_S1 Homo sapi	3.50	7.19
	307657	A1306428	Hs.298262	ribosomal protein S19	1.76	2.44
	307691	A1318285		gbxib17b01.x1 NCL_CGAP_Ov37 Homo sapiens	1.59	1.31
	307701	A1318583	Hs.276672	EST, Weakly similar to RL6_HUMAN 60S RI	1.90	2.13
	307718	A1333408	Hs.83753	small nuclear ribonucleoprotein polypept	0.45	0.99
45	307730	A1336092		gbxql43b07.x1 Soares_fetal_lung_NbHL19W	1.51	0.99
	307760	A1342387		gbxql27f07.x1 Soares_pregnant_uterus_NbH	1.00	1.00
	307764	A1342731		gbxqo26a07.x1 NCL_CGAP_Lu5 Homo sapiens	4.52	12.58
	307783	A1347274		gbxqc05d02.x1 NCL_CGAP_Co16 Homo sapiens	1.42	1.00
	307796	A1350556		gbxql18f09.x1 NCL_CGAP_GC4 Homo sapiens	6.57	9.61
50	307807	A1351799		gbxql09d02.x1 NCL_CGAP_GC4 Homo sapiens	3.38	7.68
	307808	A1351826		gbxql09g03.x1 NCL_CGAP_GC4 Homo sapiens	0.33	0.86
	307820	A1355761		gbxql94a11.x1 NCL_CGAP_Co14 Homo sapiens	7.94	21.57
	307830	A1358722	Hs.276737	EST, Weakly similar to RSHU22 ribosomal	2.05	3.32
	307852	A1365541		gbxqz08g05.x1 NCL_CGAP_CLL1 Homo sapiens	3.18	5.21
55	307902	A1380462		gbxlg02h05.x1 NCL_CGAP_CLL1 Homo sapiens	3.13	4.99
	307997	A1434512	Hs.181165	eukaryotic translation elongation factor	1.00	3.01
	308002	A1435240	Hs.283442	ESTs	5.86	12.64
	308011	A1439473		gbxib60a08.x1 NCL_CGAP_Lym12 Homo sapien	3.79	5.83
	308023	A1452732	Hs.251577	hemoglobin, alpha 1	0.38	0.88
60	308041	A1458824	Hs.169476	glyceraldehyde-3-phosphate dehydrogenase	4.36	6.05
	308059	A1468938	Hs.276877	EST, Weakly similar to RL10_HUMAN 60S R	1.80	1.98
	308085	A1474135	Hs.181165	eukaryotic translation elongation factor	3.38	4.14
	308101	A1475950	Hs.181165	eukaryotic translation elongation factor	1.30	3.87
	308106	A1476803		gbxjl77e12.x1 Soares_NSF_F8_9W_OT_PA_P_S2.38		8.72
65	308122	A1480123	Hs.309411	EST	2.70	3.86
	308154	A1500800		gbxtn93d08.x1 NCL_CGAP_UI2 Homo sapiens	0.66	1.33
	308171	A1523632	Hs.298766	ESTs, Weakly similar to schlafen4 [M.mu	2.48	4.86
	308211	A1557029	Hs.278572	anaplastic lymphoma kinase (K1-1)	2.43	2.14
	308213	A1557041		gbxPT2.1_12_E04.r tumor2 Homo sapiens cD	3.34	3.79
70	308216	A1557135		gbxPT2.1_13_H05.r tumor2 Homo sapiens cD	4.61	4.78
	308219	A1557246		gbxPT2.1_15_D07.r tumor2 Homo sapiens cD	4.87	7.94
	308271	A1567844	Hs.252259	ribosomal protein S3	2.40	6.35
	308319	A1583983	Hs.181165	eukaryotic translation elongation factor	2.45	3.33
	308362	A1613519	Hs.105749	KIAA0553 protein	1.24	1.41
75	308413	A1636253	Hs.196511	ESTs	3.16	4.82
	308450	A1650860	Hs.96840	KIAA1527 protein	1.79	2.68
	308464	A1672425	Hs.277117	EST, Moderately similar to I38055 myosi	4.87	8.27
	308588	A1718299		gbxas51g12.x1 Barstead aorta HPLRB6 Homo	3.90	5.64
	308599	A1719893		gbxas47d07.x1 Barstead aorta HPLRB6 Homo	3.32	5.12
80	308615	A1738593	Hs.101774	hypothetical protein FLJ23045	3.11	2.36
	308643	A1745040		gbxtr19a12.x1 NCL_CGAP_Ov23 Homo sapiens	3.98	3.69
	308673	A1760864		gbxwi09c10.x1 NCL_CGAP_CLL1 Homo sapiens	0.82	0.99
	308697	A1767143		gbxwi97a07.x1 NCL_CGAP_Kid12 Homo sapien	2.76	5.59
	308762	A1807405	Hs.259408	ESTs	3.17	6.30
85	308778	A1811109		gbxtr04c11.x1 NCL_CGAP_Ov23 Homo sapiens	1.00	1.00
	308782	A1811767	Hs.2186	eukaryotic translation elongation factor	2.94	5.15
	308808	A1818289		gbxwk52c01.x1 NCL_CGAP_Pr22 Homo sapiens	4.41	8.34
	308823	A1824118	Hs.217493	annexin A2	1.85	1.92
	308875	A1832332		gbxal48g03.x1 Barstead colon HPLRB7 Homo	2.52	3.80

	308879	AI832763	Hs.75968	thymosin, beta 4, X chromosome	3.38	7.96
	308886	AI833240		gb:at76d10.x1 Barstead colon HPLR87 Homo	3.06	2.65
	308898	AI858845		gb:wl32d10.x1 NCL_CGAP_U11 Homo sapiens	2.45	3.44
5	308934	AI865023	Hs.177	phosphatidylinositol glycan, class H	4.14	6.76
	308966	AI870704		gb:wl47h01.x1 NCL_CGAP_U11 Homo sapiens	1.00	1.00
	308979	AI873111		gb:wl52h05.x1 NCL_CGAP_Bm25 Homo sapien	7.15	11.10
	309045	AI910902		gb:lg39d01.x1 NCL_CGAP_U11 Homo sapiens	0.61	0.59
	309051	AI911975		gb:wd78d01.x1 NCL_CGAP_Lu24 Homo sapiens	1.78	4.42
10	309069	AI917366	Hs.78202	SWI/SNF related, matrix associated, act	3.27	5.88
	309083	AI922426	Hs.119598	ribosomal protein L3	2.39	3.34
	309105	AI925503	Hs.265884	ESTs	5.54	17.78
	309122	AI928178		gb:wo95a11.x1 NCL_CGAP_Kid11 Homo sapien	1.00	2.92
	309128	AI928816	Hs.180842	ribosomal protein L13	1.38	5.55
15	309164	AI937761		gb:wp84b09.x1 NCL_CGAP_Bm25 Homo sapien	2.43	3.11
	309177	AI951118		gb:wx63g05.x1 NCL_CGAP_Br18 Homo sapiens	0.81	0.97
	309288	AI991525	Hs.299426	ESTs	4.86	7.46
	309299	AW003478		gb:wxq66c06.x1 NCL_CGAP_GC6 Homo sapiens	4.36	9.43
	309303	AW004823		gb:ws93a08.x1 NCL_CGAP_Co3 Homo sapiens	2.88	7.54
20	309411	AW085201	Hs.244144	EST	4.30	7.14
	309437	AW090702	Hs.278242	tubulin, alpha, ubiquitous	2.49	3.11
	309459	AW117645	Hs.65114	keratin 18	2.88	4.55
	309476	AW129368		gb:xe14b05.x1 NCL_CGAP_U14 Homo sapiens	2.08	6.60
	309499	AW136325	Hs.279771	Homo sapiens clone PP1598 unknown mRNA	2.82	3.55
25	309529	AW150807	Hs.181357	laminin receptor 1 (67kD, ribosomal pro	4.78	3.95
	309532	AW151119		gb:xxg33e10.x1 NCL_CGAP_U11 Homo sapiens	1.18	4.40
	309626	AW192004	Hs.297681	serine (or cysteine) proteinase inhibiti	4.46	12.06
	309641	AW194230	Hs.253100	EST, Moderately similar to GHU Ig gamm	1.47	1.39
	309675	AW205681	Hs.253506	EST, Moderately similar to ATPN_HUMAN A	5.68	15.20
30	309693	AW237221	Hs.181357	laminin receptor 1 (67kD, ribosomal prot	1.00	1.00
	309695	AW238011	Hs.295605	mannosidase, alpha, class 2A, member 2	5.45	9.61
	309700	AW241170	Hs.179661	tubulin, beta polypeptide	1.41	1.25
	309747	AW264889		gb:xxq36h02.x1 NCL_CGAP_Lu28 Homo sapiens	5.00	8.35
35	309769	AW272346		gb:xs13c10.x1 NCL_CGAP_Kid11 Homo sapien	5.76	11.90
	309782	AW275156	Hs.156110	immunoglobulin kappa constant	0.42	0.69
	309783	AW275401	Hs.254798	EST	1.00	4.11
	309799	AW276964		gb:xxp58h01.x1 NCL_CGAP_Ov39 Homo sapiens	1.68	1.44
	309866	AW299916		gb:xs44c01.x1 NCL_CGAP_Kid11 Homo sapien	3.02	5.04
40	309903	AW339071	Hs.300697	immunoglobulin heavy constant gamma 3 (G	1.05	1.18
	309923	AW340684		gb:hd05g08.x1 Soares_NFL_T_GBC_S1 Homo s	2.30	3.67
	309928	AW341418		gb:hd08c03.x1 Soares_NFL_T_GBC_S1 Homo s	7.41	13.71
	309931	AW341683		gb:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	1.20	12.70
	309933	AW341936		gb:hb73f10.x1 NCL_CGAP_U12 Homo sapiens	4.90	18.29
	309964	AW449111	Hs.257111	hypothetical protein MGC3265	1.99	3.07
45	310002	AI439096	Hs.323079	Homo sapiens mRNA; cDNA DKFZp564P116 (tr	0.20	0.47
	310096	AW136822	Hs.172824	ESTs, Weakly similar to B48013 proline-r	1.51	1.22
	310098	AI685841	Hs.161354	ESTs	0.31	0.76
	310109	AI203094	Hs.148633	ESTs	2.06	5.83
	310112	AW197233	Hs.147253	ESTs	2.92	3.55
50	310115	AI611317	Hs.223796	ESTs	1.25	0.84
	310121	AW195642	Hs.148901	ESTs	1.00	2.71
	310146	AI206614	Hs.197422	ESTs	9.50	15.31
	310193	AI627653	Hs.147562	ESTs	2.85	4.18
	310255	AW450439	Hs.153378	ESTs	4.26	10.63
55	310261	AI240483	Hs.201217	ESTs	3.28	4.40
	310264	AI915771	Hs.74170	metallothionein 1E (functional)	0.26	0.86
	310275	AI242102	Hs.213636	ESTs	5.43	8.19
	310282	AI243332	Hs.156055	ESTs	3.15	8.06
	310290	AW013815	Hs.149103	ESTs	2.19	3.12
60	310333	AI253200	Hs.145402	ESTs	1.17	1.91
	310346	AI261340	Hs.145517	ESTs	4.81	9.95
	310385	AI263392	Hs.156151	ESTs	5.96	7.79
	310443	AW119018	Hs.164231	ESTs	2.90	4.63
	310444	AW196632	Hs.252956	ESTs	0.85	1.01
65	310446	AI275715	Hs.145926	ESTs	2.18	3.85
	310468	AI984074	Hs.196398	ESTs	3.39	5.19
	310477	AI948801	Hs.171073	ESTs	1.00	1.00
	310512	AW275603	Hs.200712	ESTs	3.87	8.12
70	310514	AI681145	Hs.160724	ESTs	3.30	7.33
	310524	AW082270	Hs.12496	ESTs, Highly similar to AC004836 1 simil	0.72	1.44
	310547	AI302654	Hs.208024	ESTs	3.26	3.46
	310584	AI653007	Hs.156304	ESTs	2.39	4.08
	310608	AI962234	Hs.196102	ESTs	5.60	6.49
	310624	AI341594		gb:Human endogenous retrovirus H proteas	4.91	9.09
75	310636	AI814373	Hs.164175	ESTs	1.85	1.71
	310648	AI347863	Hs.156672	ESTs	0.17	0.69
	310694	AI654370	Hs.157752	Homo sapiens mRNA full length insert cDN	5.40	13.22
	310695	AI472124	Hs.157757	ESTs	4.82	6.27
	310714	AI418446	Hs.157882	ESTs	1.76	3.51
80	310722	AI989803	Hs.157289	ESTs	1.14	6.85
	310755	AI916560	Hs.158707	ESTs	8.46	13.01
	310764	AI376769	Hs.167172	ESTs	4.76	7.37
	310848	AI459554	Hs.161286	ESTs	2.84	1.96
	310851	AW291714	Hs.221703	ESTs	1.00	2.32
85	310854	AI421677	Hs.161332	ESTs	6.37	7.94
	310858	AI871000	Hs.161330	ESTs	6.07	9.84

	310864	AI924558	Hs.161399	ESTs	0.87	0.78
	310875	T47764	Hs.132917	ESTs	1.00	3.63
	310896	AW157731	Hs.270982	ESTs, Moderately similar to ALU7_HUMAN A	7.07	16.68
5	310922	AW195634	Hs.170401	ESTs	1.00	1.00
	310955	AI560210	Hs.263912	ESTs	10.08	17.66
	310957	AW190974	Hs.196918	ESTs	2.18	3.18
	311000	AI521830	Hs.171050	ESTs	3.06	6.64
	311012	AW298070	Hs.241097	ESTs	1.23	3.77
10	311034	AI564023	Hs.311389	ESTs, Moderately similar to PT0375 natur	2.44	2.09
	311074	AW290922	Hs.199848	ESTs	6.04	14.19
	311134	AI900849	Hs.196971	ESTs	3.64	6.96
	311174	AW450552	Hs.205457	perixin	0.65	0.95
	311187	AI638374	Hs.224189	ESTs	2.46	2.78
15	311220	AI656040	Hs.196532	ESTs	1.10	2.52
	311230	AI989808	Hs.197663	ESTs	1.41	1.75
	311236	AI653378	Hs.197674	ESTs	2.18	2.11
	311242	AW016812	Hs.200266	ESTs	0.63	5.11
	311258	AI671221	Hs.199887	ESTs	1.00	1.41
20	311277	AW072813	Hs.270868	ESTs, Moderately similar to ALU4_HUMAN A	2.56	1.94
	311294	AA826425	Hs.291829	ESTs	1.04	2.69
	311308	F12664	Hs.49000	ESTs	1.96	6.70
	311351	AI682303	Hs.201274	ESTs	4.77	9.38
	311390	AW392987	Hs.202280	ESTs	2.80	6.06
25	311405	AW290961	Hs.201815	ESTs	3.80	11.66
	311409	AI698839		gbwd31f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.84	6.94
	311420	AI936291	Hs.209867	ESTs	5.30	12.56
	311443	AI791521	Hs.192206	ESTs	4.39	6.09
	311467	AI934909	Hs.175377	ESTs	1.00	1.04
30	311479	AI933672	Hs.211399	ESTs	2.76	5.61
	311488	R57390	Hs.301064	arfaplin 1	2.50	5.73
	311495	AW300077	Hs.221358	ESTs	3.63	6.09
	311511	AW444568	Hs.210303	ESTs	2.00	2.87
	311534	AW130351	Hs.243549	ESTs	0.31	1.33
35	311537	AI805121	Hs.211828	ESTs	3.69	5.85
	311543	AI681360	Hs.201259	ESTs	1.73	1.34
	311551	AW449774	Hs.296380	POM (POM121 rat homolog) and ZP3 fusion	3.31	6.12
	311557	AI819230	Hs.211238	interleukin-1 homolog 1	1.00	1.00
	311558	Z44432	Hs.63128	KIAA1292 protein	2.25	3.41
40	311559	AW008271	Hs.265848	similar to rat myomegalin	2.68	5.90
	311563	AI922143	Hs.211334	ESTs	2.39	3.32
	311586	AI827834	Hs.211227	ESTs	2.47	3.85
	311616	AW450575	Hs.212709	ESTs	1.00	1.00
	311621	AI924307	Hs.213464	ESTs	4.16	6.74
45	311635	AI926456	Hs.213081	ESTs	2.17	3.76
	311668	AW193574	Hs.240044	ESTs	2.60	3.12
	311672	R11807	Hs.20914	hypothetical protein FLJ123056	2.79	5.18
	311683	AW183738	Hs.232644	ESTs	0.19	0.96
50	311700	R49601	Hs.171495	retinoic acid receptor, beta	6.28	8.83
	311714	AW131785	Hs.246831	ESTs, Weakly similar to CLK_HUMAN VOLTA	5.00	8.17
	311735	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	0.98	0.72
	311743	T99079	Hs.191194	ESTs	1.00	1.96
	311783	AI682478	Hs.13528	hypothetical protein FLJ14054	0.16	0.77
	311785	AI066769	Hs.133512	ESTs	1.34	3.97
55	311799	AA780791	Hs.14014	ESTs, Weakly similar to KIAA0973 protein	8.52	13.32
	311819	AW265275	Hs.254325	ESTs	3.58	3.91
	311823	AI089422	Hs.131297	ESTs	1.40	1.72
	311877	AA349893	Hs.85339	G protein-coupled receptor 39	0.95	0.91
	311886	AA522738	Hs.132554	ESTs	0.88	0.87
60	311896	AW206447		gb:U1-H-B11-atg-g-02-0-U1.s1 NCL_CGAP_Su	1.66	1.13
	311910	N28365	Hs.22579	Homo sapiens clone CDABP0036 mRNA sequen	1.66	2.30
	311923	T60843	Hs.189579	ESTs	0.42	2.63
	311933	AI597963	Hs.118726	ESTs	1.88	3.02
	311959	T67262	Hs.124733	ESTs	2.02	2.33
65	311960	AW440133	Hs.189690	ESTs	3.87	6.62
	311967	AI382726	Hs.182434	ESTs	5.80	8.14
	311975	AA804374	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	0.98	3.26
	312005	T78450	Hs.13941	ESTs	0.12	1.39
	312028	T78886	Hs.284450	ESTs	3.78	4.92
70	312046	AI580018	Hs.268591	ESTs	4.11	7.32
	312056	T83748	Hs.268594	ESTs	2.36	3.08
	312064	AA676713	Hs.191155	ESTs	3.34	5.28
	312088	AW303760	Hs.13685	ESTs	1.60	1.15
	312093	T91809	Hs.121296	ESTs	0.68	0.85
75	312094	Z78390		gb:HSZ78390 Human fetal brain S. Meier-E	3.05	4.48
	312097	AI352096	Hs.112180	zinc finger protein 148 (pH2-52)	4.52	9.70
	312118	T85332	Hs.178294	ESTs	2.40	2.60
	312128	AI052609	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	2.39	3.53
	312147	T89855	Hs.195648	ESTs	0.67	1.03
80	312175	AA953383	Hs.127554	ESTs	5.85	10.60
	312179	AI052572	Hs.268664	ESTs	2.41	3.32
	312201	AI928365	Hs.91139	solute carrier family 1 (neuronal/epith	0.24	0.89
	312207	H90213	Hs.191330	ESTs	2.20	4.55
	312220	N74613		gb:za55a07.s1 Soares fetal liver spleen	4.28	11.13
85	312252	AI128388	Hs.143655	ESTs	1.64	1.57
	312304	AA491949	Hs.269392	ESTs	0.12	2.47



	312318	AW235092	Hs.143981	ESTs	3.46	5.69
	312319	AA216698	Hs.180780	TERA protein	5.78	4.46
	312321	R65210	Hs.186937	ESTs	0.44	1.74
5	312331	AA825512	Hs.289101	glucose regulated protein, 58kD	3.73	5.96
	312339	AA524394	Hs.165544	ESTs	3.07	0.95
	312363	AI675558	Hs.181867	ESTs	10.08	16.73
	312375	AI375096	Hs.172405	cell division cycle 27	2.78	3.71
	312376	R52089	Hs.172717	ESTs	1.00	1.00
10	312389	AI863140		gb:bt43h12.x1 NCI_CGAP_Bm52 Homo sapien	2.37	3.98
	312437	AA995028		gb:RC4-BT0629-120200-011-b10 BT0629 Homo	4.06	5.41
	312440	AI051133	Hs.133315	Homo sapiens mRNA: cDNA DKFZp761J1324 [r	1.00	1.00
	312451	R59989	Hs.176539	ESTs	4.95	10.04
	312458	AI167637	Hs.146924	ESTs	1.11	1.00
15	312507	AI168177	Hs.143653	ESTs	5.89	8.24
	312520	AI742591	Hs.205392	ESTs	3.30	8.92
	312548	AI566228	Hs.159426	hypothetical protein PRO2121	1.38	1.65
	312564	H21520	Hs.35088	ESTs	0.40	0.77
	312583	AI193122	Hs.124141	ESTs	0.13	0.94
20	312599	AI865073	Hs.125720	ESTs	3.75	5.29
	312602	AA046451	Hs.165200	ESTs	6.78	12.93
	312645	H52121	Hs.193007	ESTs	0.38	1.13
	312666	AI240582	Hs.214678	ESTs	0.98	2.03
	312689	AW450461	Hs.203965	ESTs	0.21	0.61
25	312817	H75459	Hs.233425	ESTs	1.51	0.85
	312846	AW152104	Hs.200879	ESTs	8.93	13.78
	312873	AI690071	Hs.283552	ESTs, Weakly similar to unnamed protein	4.20	6.23
	312893	AI016204	Hs.172922	ESTs	2.67	3.15
	312902	AW292797	Hs.130316	ESTs, Weakly similar to T2D3_HUMAN TRANS	1.19	0.71
30	312925	N80868	Hs.271685	ESTs	2.60	4.25
	312936	AI681581	Hs.121525	ESTs	1.00	1.17
	312975	AI640506	Hs.293119	ESTs, Weakly similar to ALU7_HUMAN ALU S	2.30	4.80
	312978	N24887	Hs.292500	ESTs	0.80	1.05
	312980	AA497043	Hs.115685	ESTs	3.12	3.60
35	312984	N25871	Hs.177337	ESTs	2.03	2.13
	313000	AI147412	Hs.145657	ESTs	5.52	8.42
	313029	AA731520	Hs.170504	ESTs	0.96	1.39
	313039	AI419290	Hs.149990	ESTs, Weakly similar to unnamed protein	6.48	13.20
	313049	AW293055	Hs.119357	ESTs	6.44	10.73
40	313056	AI651930	Hs.135684	ESTs	1.51	2.04
	313058	D81015	Hs.125382	ESTs	0.25	1.50
	313070	AI422023	Hs.161338	ESTs	8.56	11.60
	313097	AI676164	Hs.204339	ESTs	3.72	4.56
	313130	AW449171	Hs.168677	ESTs	3.28	5.06
45	313136	N59284	Hs.288010	ESTs	0.49	1.36
	313153	AI240838	Hs.132750	ESTs	5.36	5.52
	313210	N74077	Hs.197043	ESTs	0.30	0.66
	313236	AW238169	Hs.83513	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.16	8.76
	313239	W19632	Hs.124170	ESTs	1.00	3.87
50	313265	N93465	Hs.121764	ESTs, Weakly similar to testicular tekti	0.74	2.06
	313267	AI770008	Hs.129583	ESTs	0.23	1.30
	313275	AI027604	Hs.159650	ESTs	6.68	9.57
	313290	AI753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	1.34	1.07
	313292	AI362991	Hs.202121	ESTs, Weakly similar to env protein [Hs	2.00	4.32
55	313325	AI420811	Hs.127832	ESTs	1.20	2.27
	313357	AW074848	Hs.201501	ESTs	4.02	5.33
	313393	AI674685	Hs.200141	ESTs	1.36	2.84
	313399	AW376889	Hs.194097	ESTs	2.58	5.26
	313414	AI241540	Hs.132933	ESTs	6.57	15.07
60	313417	AA741151	Hs.137323	ESTs	0.63	3.01
	313457	AA576052	Hs.193223	Homo sapiens cDNA FLJ11646 fis, clone HE	2.78	4.70
	313499	AI261390	Hs.146085	KIAA1345 protein	0.91	2.37
	313516	AA029058	Hs.135145	ESTs	3.41	7.08
	313556	AA628517	Hs.118502	ESTs	0.23	0.70
65	313569	AI273419	Hs.135146	hypothetical protein FLJ13984	1.88	1.00
	313570	AA041455	Hs.209312	ESTs	0.73	2.27
	313638	AI753075	Hs.104627	Homo sapiens cDNA FLJ10158 fis, clone HE	1.00	1.72
	313662	AA740151	Hs.130425	ESTs	0.20	1.42
	313671	W49823	Hs.104613	RP42 homolog	1.00	1.00
70	313672	AW468891	Hs.122948	ESTs	3.46	5.80
	313690	AI493591	Hs.78146	platelet/endothelial cell adhesion molec	0.51	0.97
	313711	AA398070	Hs.133471	ESTs	0.18	1.01
	313723	AA070412		gb:zmm58c10.s1 Stratagene neuroepithelium	1.08	1.03
	313726	AI744687	Hs.257806	ESTs	2.13	2.99
75	313774	AW135836	Hs.144583	ESTs	1.38	1.19
	313784	AA910514	Hs.134905	ESTs	3.88	5.78
	313790	AW078569	Hs.177043	ESTs	0.22	2.06
	313832	AW271022	Hs.133294	ESTs	1.15	0.91
	313834	AW418779	Hs.114889	ESTs	0.68	3.14
80	313835	AI538438	Hs.159087	ESTs	5.74	8.88
	313852	H18633	Hs.123641	protein tyrosine phosphatase, receptor t	0.16	1.14
	313854	AW470806	Hs.275002	ESTs	2.09	4.06
	313865	AA731470	Hs.163839	ESTs	3.41	4.09
	313871	AW471088	Hs.145950	ESTs	5.28	6.83
85	313883	AI949384		gb:nu76d01.s1 NCI_CGAP_AIv1 Homo sapiens	2.90	10.91
	313915	AI969390	Hs.163443	Homo sapiens cDNA FLJ11576 fis, clone HE	1.00	1.00

	313926	AW473830	Hs.171442	ESTs	3.40	4.11
	313948	AW452823	Hs.135268	ESTs	5.77	9.15
	313978	AI870175	Hs.13957	ESTs	0.46	0.75
5	313983	AI829133	Hs.226780	ESTs	4.10	6.40
	314035	AA164199	Hs.270152	ESTs	5.88	7.90
	314037	AW300048	Hs.275272	ESTs	1.00	3.79
	314040	AA166970	Hs.118748	ESTs	7.60	11.33
	314067	AW293538	Hs.51743	KIAA1340 protein	1.86	1.21
10	314103	AI028477	Hs.132775	ESTs	2.90	5.29
	314107	AA806113	Hs.189025	ESTs	2.00	1.66
	314113	AA218985	Hs.118854	ESTs	0.91	4.17
	314124	AW118745	Hs.9460	Homo sapiens mRNA; cDNA DKFZp547C244 (fr	2.53	3.32
	314126	AA226431		gb:nc18b12.s1 NCL_CGAP_Pr1 Homo sapiens	3.13	5.08
15	314128	AA935633	Hs.194628	ESTs	2.90	6.35
	314151	AA236163	Hs.202430	ESTs	4.15	6.45
	314184	AW081795	Hs.233465	ESTs	3.44	4.65
	314192	AW290975	Hs.118923	ESTs	1.00	1.23
	314244	AL036450	Hs.103238	ESTs	2.88	3.67
20	314253	AA278679	Hs.189510	ESTs	4.98	7.16
	314262	AW086215	Hs.246096	ESTs	0.38	1.94
	314320	AA811598	Hs.275809	ESTs	3.34	5.66
	314332	AL037551	Hs.95612	ESTs	2.85	2.09
	314335	AA287443	Hs.142570	Homo sapiens clone 24629 mRNA sequence	4.35	4.78
25	314340	AW304350	Hs.130879	ESTs, Moderately similar to putative p15	0.77	0.86
	314351	AA292275	Hs.193746	ESTs	3.07	3.77
	314376	AI626633	Hs.324679	ESTs	4.10	6.11
	314443	AA827125	Hs.192043	ESTs	6.20	13.67
	314458	AI217440	Hs.143873	ESTs	0.58	2.49
30	314466	AA767818	Hs.122707	ESTs	2.53	2.62
	314478	AI521173	Hs.125507	DEAD-box protein	3.94	5.65
	314482	AL043807	Hs.134182	ESTs	1.30	1.44
	314506	AA833655	Hs.206868	Homo sapiens cDNA FLJ14056 fis, clone HE	3.28	3.47
	314519	R42554	Hs.210862	T-box, brain, 1	3.12	6.16
35	314529	AL046412	Hs.202151	ESTs	3.43	6.87
	314546	AW007211	Hs.16131	hypothetical protein FLJ12876	1.38	1.00
	314552	AI564127	Hs.143493	ESTs	2.29	5.27
	314579	AW197442	Hs.116998	ESTs	3.87	5.75
	314580	AW451832	Hs.255938	ESTs, Moderately similar to KIAA1200 pro	0.10	0.71
40	314585	AA918474	Hs.216363	ESTs	1.08	1.40
	314589	AW384790	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	1.00	1.00
	314592	AA435781	Hs.192148	ESTs	0.90	2.60
	314603	AA418024	Hs.270570	ESTs	4.56	6.29
	314604	AA946582	Hs.8700	deleted in liver cancer 1	3.42	3.92
45	314606	AA418241	Hs.188767	ESTs	2.97	4.55
	314648	AA878419		gb:EST391378 MAGE resequences, MAGP Homo1.42		1.36
	314699	AI038719	Hs.132801	ESTs	3.66	4.97
	314701	AI754634	Hs.131987	ESTs	0.03	0.90
	314710	AI669131	Hs.290989	EST	3.40	7.52
50	314750	AI095005	Hs.135174	ESTs	2.80	6.54
	314767	AW135412	Hs.164002	ESTs	3.20	4.26
	314801	AA481027	Hs.109045	hypothetical protein FLJ10498	1.00	1.00
	314817	AI694139	Hs.192855	ESTs	0.91	0.99
	314835	AI281370	Hs.76084	ribosomal protein L27a	5.75	7.44
55	314852	AI903735		gb:MR-BT035-200199-031 BT035 Homo sapien	1.68	4.34
	314853	AA729232	Hs.153279	ESTs	0.60	1.85
	314940	AW452768	Hs.162045	ESTs	10.10	16.20
	314941	AA515902	Hs.130650	ESTs	0.31	1.02
	314943	AI476797	Hs.184572	cell division cycle 2, G1 to S and G2 to	2.18	0.37
60	314955	AA521382	Hs.192534	ESTs	2.59	3.90
	314973	AW273128	Hs.300268	ESTs	1.05	1.25
	315004	AA527941	Hs.325351	EST	5.64	13.63
	315006	AI538613	Hs.298241	Transmembrane protease, serine 3	0.52	1.78
	315033	AI493046	Hs.146133	ESTs	2.46	1.00
65	315035	AI569476	Hs.177135	ESTs	0.34	1.33
	315056	AI202703	Hs.152414	ESTs	2.10	2.64
	315069	AI821517	Hs.105866	ESTs	1.00	1.30
	315071	AA552690	Hs.152423	Homo sapiens cDNA: FLJ121274 fis, clone C	1.78	1.00
70	315073	AW452948	Hs.257631	ESTs	1.17	1.52
	315078	AA568548	Hs.190616	ESTs	3.00	3.79
	315080	AA744550	Hs.136345	ESTs	1.00	1.00
	315120	AA564991	Hs.269477	ESTs	0.64	1.44
	315175	AI025842	Hs.152530	ESTs	0.61	1.91
	315193	AI241331	Hs.131765	ESTs	1.05	0.97
75	315196	AA972756	Hs.44898	Homo sapiens clone TCCCTA00151 mRNA sequ	0.48	1.96
	315200	AI808235	Hs.307686	EST	3.76	9.40
	315254	AI474433	Hs.179556	ESTs	5.37	9.36
	315353	AW452608	Hs.279610	hypothetical protein FLJ10493	1.00	1.30
	315397	AA218940	Hs.137516	fidgulin-like 1	3.38	2.24
80	315403	AW362880	Hs.163924	ESTs	2.04	5.23
	315431	AA622104	Hs.184838	ESTs	2.36	8.04
	315454	AI239473		gb:zh36f02.x1 Soares_NFL_T_GBC_S1 Homo s	3.46	7.64
	315455	AW393391	Hs.156919	ESTs	3.78	5.76
	315473	AI691671	Hs.312671	ESTs, Moderately similar to OVCA1	0.89	2.15
85	315483	AW512763	Hs.222024	transcription factor BMAL2	2.32	1.96
	315526	AI193048	Hs.128685	ESTs	1.67	1.78

	315530	AI200852	Hs.127780	ESTs	1.05	1.01
	315541	AI168233	Hs.123159	sperm associated antigen 4	0.85	0.56
	315552	AW445034	Hs.255578	ESTs	1.00	2.22
	315562	AA737415	Hs.152826	ESTs	2.66	2.48
5	315577	AW513545	Hs.17283	hypothetical protein FLJ10890	2.20	2.25
	315587	AI268399	Hs.140489	ESTs	1.00	1.04
	315589	AW072387	Hs.155258	Homo sapiens mRNA; cDNA DKFZp434B1272 (f	0.14	1.05
	315623	AA354078	Hs.258189	ESTs	7.44	12.56
10	315634	AA837085	Hs.220585	ESTs	0.50	1.40
	315668	AA912347	Hs.136585	ESTs	0.43	1.22
	316677	AI932662	Hs.164073	ESTs	0.60	1.39
	315706	AW440742	Hs.155556	hypothetical protein FLJ20202	2.18	3.77
	315707	AI418055	Hs.161160	ESTs	2.88	2.63
	315730	H25899	Hs.201591	ESTs	0.11	0.60
15	315745	AI821759	Hs.191856	ESTs	3.50	7.25
	315791	AA678177		gbz115a05.s1 Soares_fetal_liver_spleen_	1.78	2.63
	315801	AA827752	Hs.266134	ESTs	4.31	6.23
	315820	AI652022	Hs.258785	ESTs	2.35	3.01
	315878	AA683336	Hs.189046	ESTs	2.12	2.64
20	315905	AI821911	Hs.209452	ESTs	1.03	1.97
	315923	AI052789	Hs.133263	ESTs	2.63	5.06
	315954	AW276810	Hs.254859	ESTs, Moderately similar to ALU5_HUMAN A	1.21	0.85
	315978	AA830893	Hs.119769	ESTs	3.09	3.41
	316001	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	2.20	6.82
25	316011	AW516953	Hs.201372	ESTs	0.35	1.63
	316012	AA764950	Hs.119898	ESTs	6.56	8.13
	316040	AI983409	Hs.189226	ESTs	5.69	10.69
	316048	AI720759	Hs.224971	ESTs	2.84	10.45
30	316076	AW297895	Hs.116424	ESTs	0.30	1.05
	316124	AI308862	Hs.167028	ESTs	1.00	1.43
	316151	AI806016	Hs.156520	ESTs	5.80	9.03
	316187	AW518299	Hs.192253	ESTs	1.20	3.96
	316204	AA731509	Hs.120257	ESTs	4.92	6.94
35	316232	AW297853	Hs.251203	ESTs	1.48	1.60
	316275	AI671041	Hs.292611	ESTs, Moderately similar to ALU1_HUMAN A	5.86	12.14
	316291	AW375974	Hs.156704	ESTs	2.73	2.69
	316303	AA740594	Hs.209609	ESTs	1.53	1.26
	316344	AA744518	Hs.120610	ESTs	3.66	8.34
40	316346	AI028478	Hs.157447	ESTs	3.51	6.69
	316365	AI627845	Hs.210776	ESTs	2.50	4.33
	316380	AI393378	Hs.164496	ESTs	1.16	2.16
	316470	AA809902	Hs.243813	ESTs	5.40	10.34
	316509	AA767310	Hs.291766	ESTs	2.46	2.89
45	316514	AA768037	Hs.291671	ESTs	4.70	6.04
	316519	AI929097		gbz110c11.s1 NCI_CGAP_GCB1 Homo sapiens	4.41	9.70
	316609	AW292520	Hs.122082	ESTs	1.00	2.89
	316633	AI125586	Hs.127955	ESTs	2.61	3.72
	316700	AW172316	Hs.252961	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.46	4.64
50	316711	AI743721	Hs.285316	ESTs, Moderately similar to ALU7_HUMAN A	4.45	6.95
	316713	AI090671	Hs.134807	hypothetical protein FLJ12057	0.30	2.40
	316715	AI440266	Hs.170673	ESTs, Weakly similar to AF126780 1 retin	0.20	1.45
	316787	AW369770	Hs.130351	ESTs	4.05	5.53
	316809	AA825839	Hs.202238	ESTs	2.25	3.82
55	316811	AA922060	Hs.132471	ESTs	1.00	1.32
	316812	AW135045	Hs.232001	ESTs	3.28	4.70
	316818	AA827176	Hs.124316	ESTs	0.67	1.81
	316824	AA837416	Hs.124299	ESTs	3.53	6.00
	316827	AI380429	Hs.172445	ESTs	0.72	1.56
60	316891	AW298119	Hs.202536	ESTs	1.64	2.97
	316951	AA134365	Hs.57548	ESTs	1.45	1.08
	316970	AA860172	Hs.132406	ESTs	1.00	1.53
	316971	AA860212	Hs.170991	ESTs	1.08	1.96
	316990	AA861611	Hs.130643	ESTs	5.44	10.04
65	317001	AI627917	Hs.233694	hypothetical protein FLJ11350	3.56	4.37
	317008	AW051597	Hs.143707	ESTs	0.69	1.37
	317051	AA873253	Hs.126233	ESTs	6.18	12.72
	317128	AA971374	Hs.125674	ESTs	1.87	2.66
	317129	H12523	Hs.78521	Homo sapiens cDNA: FLJ21193 fis, clone C	4.12	6.84
70	317137	AW341567	Hs.125710	ESTs	2.82	5.12
	317198	AI348258	Hs.153412	ESTs	1.98	2.51
	317212	AI866468	Hs.148294	ESTs	1.88	2.83
	317223	AW297920	Hs.130054	ESTs	0.83	1.57
	317224	D56760	Hs.93029	sparc/osteonectin, cwcw and kazal-like d	2.74	0.86
75	317266	AA906289	Hs.203614	ESTs	1.00	1.00
	317282	AI807444	Hs.176101	ESTs	2.60	4.21
	317285	AW370682	Hs.222080	ESTs	1.96	3.49
	317302	AA908709	Hs.135564	ESTs	7.16	8.32
	317304	AW449899	Hs.130184	ESTs	1.38	2.28
80	317320	AA927151	Hs.130452	ESTs	3.58	8.13
	317413	AW341701	Hs.126622	ESTs	2.08	4.92
	317417	AA918420	Hs.145378	ESTs	3.06	4.79
	317452	AA972955	Hs.135568	ESTs	4.22	9.21
	317519	AI859895	Hs.126860	ESTs	1.88	4.15
85	317521	AI824338	Hs.126891	ESTs	3.12	4.55
	317529	AI916517	Hs.126865	ESTs	2.73	3.34

	317570	AI733361	Hs.127122	ESTs	1.00	2.43
	317571	AA938663	Hs.199828	ESTs	5.20	11.95
	317598	AW206035	Hs.192123	ESTs	0.33	1.56
5	317627	AI346110	Hs.132553	ESTs	1.50	1.39
	317650	AI733310	Hs.127346	ESTs	0.48	1.46
	317659	AA961216	Hs.127785	ESTs	4.18	7.14
	317674	AW294909	Hs.132208	ESTs	2.92	3.20
	317686	AA969051	Hs.187319	ESTs	1.00	1.01
10	317692	AI307659	Hs.174794	ESTs	5.33	9.59
	317701	AI674774	Hs.128014	ESTs	1.00	1.00
	317711	AI733015	Hs.272189	ESTs	5.13	7.81
	317722	AI733373	Hs.128119	ESTs	2.50	6.03
	317756	AA973667	Hs.128320	ESTs	1.59	1.30
15	317777	AI143525	Hs.47313	KIAA0258 gene product	1.00	2.48
	317799	AI498273	Hs.128808	ESTs	1.78	2.11
	317803	AA983251	Hs.128899	ESTs	0.80	1.08
	317821	AI368158	Hs.70983	PTPL1-associated RhoGAP 1	0.17	0.68
	317848	AI820575	Hs.129086	Homo sapiens cDNA FLJ12007 fis, clone HE	5.30	8.16
20	317850	N29974	Hs.152982	hypothetical protein FLJ13117	1.30	2.28
	317861	AW341064	Hs.129119	ESTs	2.18	5.93
	317865	AI298794	Hs.129130	ESTs	4.48	8.20
	317869	AW295184	Hs.129142	deoxyribonuclease II beta	0.44	0.99
	317881	AI827248	Hs.224398	Homo sapiens cDNA FLJ11469 fis, clone HE	4.06	2.23
25	317890	AI915599	Hs.128225	ESTs	4.68	7.48
	317899	AI952430	Hs.150614	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.14	3.37
	317986	AI005163	Hs.201378	ESTs, Weakly similar to T12545 hypothetical	0.28	1.66
	318001	AW235697	Hs.130980	ESTs	5.12	9.97
	318016	AI016694	Hs.256921	ESTs	1.86	4.50
30	318023	AW243058	Hs.131155	ESTs	2.92	5.22
	318054	AW449270	Hs.232140	ESTs	3.92	6.37
	318068	AI024540	Hs.131574	ESTs	1.21	1.27
	318117	AI208304	Hs.250114	ESTs	0.86	1.17
	318187	AI792585	Hs.133272	ESTs, Weakly similar to ALUC_HUMAN IIII	5.90	6.98
35	318223	AI077540	Hs.134090	ESTs	1.05	0.90
	318240	AI085377	Hs.143610	ESTs	3.10	2.40
	318255	AI082692	Hs.134662	ESTs	0.02	1.05
	318266	AI554341	Hs.271443	ESTs	6.12	10.55
	318330	AI093840	Hs.143758	ESTs	4.98	7.90
40	318369	AI493501	Hs.170974	ESTs	2.46	5.62
	318428	AI949409	Hs.194591	ESTs	0.77	0.45
	318458	AI149783	Hs.158438	ESTs	3.54	4.92
	318467	AI151395	Hs.144834	ESTs	4.56	5.62
	318473	AI939339	Hs.146883	ESTs	2.08	4.05
45	318476	AI693927	Hs.265165	ESTs	4.22	8.07
	318487	AI157877	Hs.143716	ESTs	1.47	1.05
	318488	AI217431	Hs.144709	ESTs	1.40	4.14
	318491	T26477	Hs.22883	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.84	1.80
	318499	T25451		gb:PTH1188 HTCDL1 Homo sapiens cDNA 5/3	2.58	5.20
50	318537	AA377908	Hs.13254	ESTs	3.26	4.18
	318538	N28625	Hs.74034	Homo sapiens clone 24651 mRNA sequence	0.35	1.07
	318547	R20578	Hs.90431	ESTs	3.22	4.60
	318552	R18364	Hs.90363	ESTs	4.87	9.05
	318575	R55102	Hs.107761	ESTs, Weakly similar to unnamed protein	1.91	1.98
55	318580	T34571	Hs.49007	poly(A) polymerase alpha	2.74	6.22
	318587	AA779704	Hs.168830	Homo sapiens cDNA FLJ12136 fis, clone MA	0.85	2.46
	318596	AI470235	Hs.172698	EST	4.88	4.93
	318622	T48325	Hs.237658	apolipoprotein A-II	4.80	12.51
	318629	N25163	Hs.8861	ESTs	0.39	1.04
60	318637	AA243539	Hs.9196	hypothetical protein	1.72	3.57
	318648	T77141	Hs.184411	albumin	6.27	9.91
	318650	AA393302	Hs.176626	hypothetical protein EDAG-1	3.96	8.84
	318671	AA188823	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	1.53	0.81
	318679	T58115	Hs.10336	ESTs	1.00	2.19
65	318711	AI936475	Hs.101282	Homo sapiens cDNA: FLJ21238 fis, clone C	3.05	3.18
	318725	AI962487	Hs.242990	ESTs	1.08	2.46
	318728	Z30201	Hs.291289	ESTs, Weakly similar to ALU1_HUMAN ALU S	0.77	1.33
	318740	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.25	1.49
	318776	R24963	Hs.23766	ESTs	1.00	3.01
70	318784	H00148	Hs.5181	proliferation-associated 2G4, 38kD	2.70	3.86
	318816	F07873	Hs.21273	ESTs	3.90	7.13
	318865	H10818		gb:ym04f10.r1 Soares Infant brain 1N1B H	2.25	3.56
	318879	R56332	Hs.18268	adenylate kinase 5	1.78	5.00
	318881	Z43224	Hs.124952	ESTs	4.79	14.13
75	318894	F08138	Hs.7387	DKFZP564B116 protein	5.31	7.00
	318901	AW368520	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	1.03	0.91
	318925	Z43577	Hs.21470	ESTs	2.23	3.80
	318936	AI219221	Hs.308298	ESTs	1.86	7.16
	318982	Z44140	Hs.269622	ESTs	5.84	9.79
80	318986	Z44186	Hs.169161	ESTs, Highly similar to MAON_HUMAN NADP-	1.00	1.00
	319041	Z44720	Hs.98385	ESTs, Weakly similar to weak similarity	3.38	6.11
	319103	H05896	Hs.4993	KIAA1313 protein	1.00	1.07
	319170	R13678	Hs.285306	putative selenocysteine lyase	3.79	5.03
	319196	F07953	Hs.16085	putative G-protein coupled receptor	1.00	2.98
85	319199	F07361	Hs.13306	ESTs	3.53	5.66
	319242	F11472	Hs.12839	ESTs	5.87	7.26

5	319263	T65331	Hs.81360	Homo sapiens cDNA: FLJ21927 fis, clone H	1.81	1.57
	319267	F11802	Hs.6818	ESTs	1.10	4.72
	319270	R13474	Hs.290263	ESTs	4.80	10.40
	319279	T65094	Hs.12677	OGI-147 protein	1.50	2.11
	319282	AA461358	Hs.12876	ESTs	1.00	1.00
10	319289	W07304	Hs.79059	transforming growth factor, beta recepto	0.18	0.68
	319291	W86578	Hs.285243	hypothetical protein FLJ22029	0.26	0.62
	319293	F12119	Hs.12583	ESTs	3.13	4.50
	319312	Z45481		gbtHSC2QE041 normalized infant brain cDN	1.10	1.00
	319370	H54254	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	0.16	0.73
15	319391	R06304	Hs.13911	ESTs	1.26	2.43
	319398	H67130	Hs.301743	ESTs	0.70	0.76
	319398	AA359754	Hs.191196	ESTs	2.45	3.59
	319407	R05329		gbys91b04.r1 Soares fetal liver spleen	2.00	3.54
	319425	T82930		gbys39f07.r1 Soares fetal liver spleen	4.28	8.81
20	319433	R06050	Hs.191198	ESTs	6.15	14.13
	319437	AA282420	Hs.111991	ESTs, Weakly similar to Y48A5A.1 [C.eleg	3.26	5.68
	319466	AI809937	Hs.116417	ESTs	1.76	5.65
	319471	R06546	Hs.19717	ESTs	4.29	4.84
	319480	R06933	Hs.184221	ESTs	1.00	1.00
25	319484	T91772		gbys52a10.s1 Soares fetal liver spleen	2.81	4.88
	319486	AI382429	Hs.250799	ESTs	2.08	2.82
	319508	T98988	Hs.270104	ESTs, Moderately similar to ALLJ8_HUMAN A	2.80	4.39
	319523	T69499	Hs.191184	ESTs	1.55	3.25
	319545	R83716	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	1.65	1.19
30	319546	R09592		gbysf23b12.r1 Soares fetal liver spleen	5.11	8.54
	319552	AA096106	Hs.20403	ESTs	1.89	3.36
	319582	T82998	Hs.250154	hypothetical protein FLJ12973	3.48	4.82
	319586	D78808	Hs.283683	chromosome 8 open reading frame 4	0.26	0.82
	319604	R11679	Hs.297753	vimentin	1.68	3.41
35	319609	AW247514	Hs.12293	hypothetical protein FLJ21103	3.06	4.24
	319611	H14957		gbym19c10.r1 Soares infant brain 1NIB H	2.76	4.24
	319653	AA770183	Hs.173515	uncharacterized hypothalamus protein HT0	2.51	3.55
	319657	R19897	Hs.106604	ESTs	5.32	7.68
	319658	R13432	Hs.167481	syntrophin, gamma 1	3.35	5.00
40	319661	H08035	Hs.21398	ESTs, Moderately similar to A Chain A, H	5.18	12.55
	319662	H06382	Hs.21400	ESTs	1.58	1.56
	319708	R15372	Hs.22664	ESTs	1.00	1.22
	319742	T77668	Hs.21162	ESTs	2.48	3.13
	319748	R18178	Hs.296866	Homo sapiens mRNA; cDNA DKFZp434N1923 (f	3.02	4.85
45	319772	R76633	Hs.22646	ESTs	4.36	11.61
	319788	AA321932	Hs.117414	KIAA1320 protein	2.56	3.68
	319805	R92857	Hs.271350	likely ortholog of mouse polydom	4.63	6.56
	319812	N74880	Hs.264330	N-acylsphingosine amidohydrolase (acid c	0.63	1.32
	319834	AA071267		gbczm61g01.r1 Stratagene fibroblast (937	0.30	0.94
50	319878	T78517	Hs.13941	ESTs	3.99	6.44
	319882	AA258981	Hs.291392	ESTs	5.09	7.36
	319912	T77559	Hs.94109	Homo sapiens cDNA FLJ13634 fis, clone PL	3.24	3.21
	319935	H79460	Hs.271722	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.40	9.42
	319944	T79248	Hs.133510	ESTs	3.31	5.39
55	319947	AA150967	Hs.14479	Homo sapiens cDNA FLJ14199 fis, clone NT	2.90	4.95
	319962	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	1.81	1.57
	320007	AA336314		gbtEST40943 Endometrial tumor Homo sapie	3.42	6.29
	320018	T83263		gbysd40h09.r1 Soares fetal liver spleen	2.77	5.14
	320030	H63789	Hs.296288	ESTs, Weakly similar to KIAA0638 protein	4.10	6.69
60	320032	AI699772	Hs.292664	ESTs, Weakly similar to A46010 X-linked	3.27	3.27
	320040	AA233671	Hs.87164	hypothetical protein FLJ14001	1.81	1.64
	320047	T86564	Hs.302256	EST	3.38	7.36
	320063	AA074108	Hs.120844	FOXJ2 forkhead factor	5.90	16.73
	320096	H58138	Hs.117915	ESTs	2.08	4.47
65	320099	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	1.00	1.00
	320112	T92107	Hs.188489	ESTs	2.27	2.06
	320140	H94179	Hs.119023	SMC2 (structural maintenance of chromoso	1.00	1.00
	320188	AW419200	Hs.172318	ESTs	1.26	1.00
	320193	AA831259	Hs.17132	ESTs	2.58	6.23
70	320195	R62203	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	2.85	4.53
	320199	R78659	Hs.29792	ESTs	0.40	0.94
	320203	AL049227	Hs.124776	Homo sapiens mRNA; cDNA DKFZp564N1116 (f	0.84	1.18
	320219	AA327564	Hs.127011	tubulointerstitial nephritis antigen	1.00	1.17
	320220	AF054910	Hs.127111	takin 2 (testicular)	0.18	1.09
75	320225	AF058989	Hs.128231	G antigen, family B, 1 (prostate associa	5.26	13.75
	320231	H03139	Hs.24683	ESTs	1.59	1.93
	320260	NM_003608	Hs.131924	G protein-coupled receptor 65	1.38	4.56
	320267	AL049337	Hs.132571	Homo sapiens mRNA; cDNA DKFZp564P016 (fr	1.00	1.92
	320268	H06019	Hs.151293	Homo sapiens cDNA FLJ10664 fis, clone NT	5.58	5.70
80	320322	AF077374	Hs.139322	small proline-rich protein 3	1.41	1.01
	320325	AI167978	Hs.139851	caveolin 2	0.05	0.67
	320330	AF026004	Hs.141660	chloride channel 2	2.17	1.26
	320339	H10807	Hs.281434	Homo sapiens cDNA FLJ14028 fis, clone HE	1.81	2.32
	320388	H16065	Hs.31286	ESTs	1.00	3.22
85	320402	R22291	Hs.23358	Homo sapiens clone FLC0578 PRO2852 mRNA,	1.41	1.36
	320413	AA203711	Hs.173269	ESTs	2.31	3.61
	320432	R62786	Hs.124136	ESTs	11.25	20.78
	320438	AA253352	Hs.293663	ESTs	2.22	3.49
	320438	W24548	Hs.5669	ESTs	3.53	8.14

	320448	A1240233	Hs.80887	v-yes-1 Yamaguchi sarcoma viral related	1.42	3.46
	320451	R26944	Hs.180777	Homo sapiens mRNA; cDNA DKFZp564M0264 (f	0.87	0.81
	320484	AA094436	Hs.296267	folliculin-like 1	0.65	1.18
5	320499	R32555	Hs.24321	Homo sapiens cDNA FLJ12028 fis, clone HE	3.44	7.15
	320514	AB007978	Hs.158278	KIAA0509 protein	6.44	13.62
	320521	N31464	Hs.24743	hypothetical protein FLJ20171	1.48	1.04
	320526	AW374205	Hs.111314	ESTs	3.66	7.87
	320527	R34672	Hs.324522	ESTs	3.16	5.63
10	320536	AA331732	Hs.137224	ESTs	2.83	5.83
	320556	AF054177	Hs.14570	hypothetical protein FLJ22530	1.28	1.00
	320564	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.22	0.81
	320587	Z44524	Hs.167456	Homo sapiens mRNA full length insert cDN	1.84	2.44
	320635	R54159	Hs.80506	small nuclear ribonucleoprotein polypept	1.00	6.25
	320639	AA243258	Hs.7395	hypothetical protein FLJ23182	2.60	2.30
15	320648	N48521	Hs.26549	Homo sapiens mRNA for KIAA1708 protein,	1.00	1.53
	320651	AA489268	Hs.111334	ferritin, light polypeptide	0.14	0.79
	320664	AI904216	Hs.91251	hypothetical protein FLJ11198	5.02	8.84
	320676	AA132650	Hs.300511	ESTs	3.63	5.37
	320683	R59291	Hs.26638	ESTs, Weakly similar to unnamed protein	0.37	1.31
20	320689	AA334609	Hs.171929	ESTs, Weakly similar to A54849 collagen	1.27	1.02
	320696	AW135016	Hs.172780	ESTs	3.53	4.60
	320714	AI445591	Hs.181125	gbcyq04a10.r1 Soares fetal liver spleen	1.06	0.85
	320727	U96044	Hs.117176	immunoglobulin lambda locus	1.35	1.49
	320771	AI793266	Hs.91226	poly(A)-binding protein, nuclear 1	0.04	0.82
25	320794	AA281993	Hs.194679	ESTs	2.96	4.33
	320822	AF100780	Hs.194689	WNT1 inducible signaling pathway protein	0.10	0.79
	320824	AF120274	Hs.266416	artemin	1.16	1.11
	320830	AJ132445	Hs.34744	claudin 14	1.06	1.75
	320843	AA317372	Hs.34771	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	1.36	1.47
30	320849	D60031	Hs.135904	ESTs	5.30	7.49
	320853	AI473796	Hs.271580	ESTs	1.00	1.00
	320896	AB002155	Hs.199538	uroplakin 1B	5.90	2.55
	320921	R94038	Hs.213923	inhibin, beta C	2.20	1.17
	320927	AI205786	Hs.92023	ESTs	0.18	1.46
35	320957	AI878933	Hs.22644	core histone macroH2A.2	1.67	2.18
	320997	H22544	Hs.293650	gbcy69f11.r1 Soares adult brain N2b5HB5	3.26	3.62
	321045	W88483	Hs.269055	ESTs	2.25	4.55
	321046	H27794	Hs.240770	ESTs	2.69	4.25
40	321052	AW372884	Hs.126465	nuclear cap binding protein subunit 2, 2	2.14	2.56
	321059	AI092824	Hs.241411	ESTs	1.69	0.53
	321062	R87955	Hs.241438	Homo sapiens mRNA full length insert cDN	2.76	5.20
	321067	AF131782	Hs.241438	Homo sapiens clone 24941 mRNA sequence	4.79	7.41
	321102	AA018306	Hs.125494	gbzye40d08.r1 Soares retina N2b4HR Homo	1.79	4.27
45	321130	H43750	Hs.298351	ESTs	1.00	3.14
	321142	AI817933	Hs.99598	ASPL protein	8.73	15.36
	321155	AA336635	Hs.172982	hypothetical protein MGC5338	3.04	5.03
	321158	AA700289	Hs.226469	gbzyu76f11.r1 Soares fetal liver spleen	4.62	8.39
	321170	N53742	Hs.226414	ESTs	2.21	4.46
50	321199	AW385512	Hs.18192	gbzyy56d10.s1 Soares multiple sclerosis	5.69	8.01
	321206	H54178	Hs.222024	Homo sapiens cDNA FLJ12417 fis, clone MA	4.00	7.32
	321225	AL080073	Hs.271770	Homo sapiens mRNA; cDNA DKFZp564B1462 (f	4.17	4.63
	321236	AW371941	Hs.108551	Ser/Arg-related nuclear matrix protein (	1.00	1.00
	321244	AF068654	Hs.159235	gbtHomo sapiens isolate AN.1 immunoglobu	2.18	9.13
55	321270	R83560	Hs.266514	gbzyv76c06.s1 Soares fetal liver spleen	3.80	5.26
	321317	AI937060	Hs.137507	KIAA1151 protein	1.81	1.55
	321318	AB033041	Hs.300646	KIAA1215 protein	1.00	1.00
	321325	AB033100	Hs.222024	KIAA protein (similar to mouse paladin)	0.44	0.93
	321342	AA127984	Hs.271770	transcription factor BMAL2	4.94	4.93
60	321356	R93443	Hs.161075	ESTs	3.10	4.66
	321418	AI739161	Hs.132743	ESTs	2.28	2.54
	321420	AI368667	Hs.82845	ESTs	1.13	0.97
	321430	U05890	Hs.39540	gbtHsapiens (DIG3) mRNA for immunoglobu	2.42	3.35
	321453	N50080	Hs.292549	Homo sapiens cDNA: FLJ21930 fis, clone H	1.60	3.11
65	321467	X13075	Hs.255436	gbtHuman 2a12 mRNA for kappa-immunoglobu	0.42	0.72
	321468	AA514198	Hs.268980	ESTs	2.46	6.50
	321491	H70665	Hs.225748	ESTs	1.00	1.25
	321498	AW295517	Hs.108551	ESTs	3.19	6.24
	321504	W02355	Hs.255748	ESTs	2.28	3.86
70	321510	AA703650	Hs.108551	ESTs	2.14	3.94
	321513	H84972	Hs.159235	ESTs	2.78	5.37
	321516	AI382803	Hs.266514	ESTs	3.06	7.19
	321565	AI525773	Hs.266514	hypothetical protein FLJ11342	4.89	7.82
	321577	H84260	Hs.28803	gbzy90g04.r1 Soares retina N2b5HR Homo	1.00	1.73
75	321581	AA019964	Hs.21858	ESTs	4.88	6.73
	321582	AA143755	Hs.96322	trinucleotide repeat containing 3	1.00	2.08
	321587	H95531	Hs.161051	gbzy76e02.r1 Soares retina N2b4HR Homo	2.26	4.52
	321626	AA295430	Hs.247084	hypothetical protein FLJ23560	1.95	3.83
	321628	H87064	Hs.294110	ESTs, Moderately similar to ALU6_HUMAN A	0.47	1.02
80	321642	AW085917	Hs.123158	ESTs	1.52	1.38
	321669	H95404	Hs.167260	ESTs	2.17	2.45
	321687	AA625149	Hs.42568	gbt:af70c12.r1 Soares_NhHMPu_S1 Homo sapi	4.31	6.95
	321688	H97646	Hs.173737	Homo sapiens cDNA FLJ12830 fis, clone NT	2.82	3.28
	321693	AA700017	Hs.167260	ras-related C3 botulinum toxin substrate	0.51	1.08
85	321700	N55160	Hs.42568	ESTs	4.57	7.46
	321701	AW390923	Hs.42568	ESTs	1.00	1.00

	321709	N25847	Hs.1028923	RAB38, member RAS oncogene family	1.00	1.00
	321710	N35682	Hs.259743	ESTs	2.97	5.26
	321775	AI694875	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca	1.00	1.00
5	321777	AI637993	Hs.202312	Homo sapiens clone N11 Ntera2D1 teratoca	1.68	0.45
	321779	N42729	Hs.163835	ESTs	0.90	0.90
	321829	D81993	Hs.8966	tumor endothelial marker 8	2.69	3.89
	321846	AA281594	Hs.87902	ESTs	5.11	7.64
	321879	AL109670	Hs.302809	ESTs	6.49	9.58
10	321883	AA426494	Hs.46901	KIAA1462 protein	0.28	0.95
	321899	N55158	Hs.29468	ESTs	0.39	0.95
	321911	AF026944	Hs.293797	ESTs	6.20	10.76
	321949	R49202	Hs.181694	EST	4.62	10.51
	321955	AI651866	Hs.195689	ESTs	2.89	5.47
	321956	AL110177	Hs.132882	ESTs	0.32	1.25
15	321987	AL133612	Hs.272759	KIAA1457 protein	1.00	1.83
	321991	AL133627	Hs.158923	Homo sapiens mRNA; cDNA DKFZp434K0722 (f	4.00	6.47
	322002	AA328801	Hs.84522	ESTs	2.10	3.48
	322035	AL137517	Hs.306201	hypothetical protein DKFZp564O1278	1.00	1.90
20	322044	AW340926		gb:xy51b10.x1 NCL_CGAP_Lu34.1 Homo sapie	3.20	9.67
	322057	N92197	Hs.154679	synaptotagmin 1	1.55	1.07
	322060	AI341937		gb:qt10e03.x1 NCL_CGAP_GC4 Homo sapiens	4.59	7.68
	322070	U80769	Hs.210322	Homo sapiens mRNA for KIAA1766 protein,	2.78	4.52
	322083	AF074882	Hs.226031	ESTs, Highly similar to KIAA0635 protein	3.10	5.52
	322091	AI819863	Hs.106243	ESTs	1.69	1.75
25	322125	R83901		gb:yy16c12.r1 Soares fetal liver spleen	2.06	5.27
	322130	R98978	Hs.117767	ESTs	10.12	16.49
	322147	AF085919	Hs.114176	ESTs	0.94	0.64
	322166	AF085958		gb:yr68b03.r1 Soares fetal liver spleen	4.09	6.67
30	322173	H52567		gb:y185d04.r1 Soares_pineal_gland_N3HPG	3.46	4.85
	322178	H56535		gb:y188g03.r1 Soares_pineal_gland_N3HPG	0.44	2.54
	322179	H92891		gb:y194c02.s1 Soares_pineal_gland_N3HPG	4.52	7.50
	322186	H67346	Hs.269187	ESTs	0.15	0.98
	322196	W87895	Hs.211516	ESTs	2.20	5.04
35	322212	AF087995	Hs.134877	ESTs	3.42	4.84
	322221	AI890619	Hs.179562	nucleosome assembly protein 1-like 1	0.82	2.14
	322277	AI640193	Hs.226389	ESTs	3.62	3.98
	322278	AF086283		gb:zd46f01.r1 Soares_fetal_heart_NbHH19W	1.00	1.00
	322284	AI792140	Hs.49265	ESTs	0.66	2.76
40	322288	AL037273	Hs.7886	pellino (Drosophila) homolog 1	0.71	0.70
	322320	AF086419		gb:zd78d03.r1 Soares_fetal_heart_NbHH19W	2.02	2.76
	322336	AA308526	Hs.76152	decorin	2.92	4.44
	322339	W17348		gb:zb18c07.x5 Soares_fetal_Jung_NbHL19W	8.50	11.56
	322366	AW404274	Hs.122492	hypothetical protein	0.61	1.34
45	322372	W25624	Hs.153943	ESTs	7.37	12.07
	322374	AI394663	Hs.122116	ESTs, Moderately similar to Ostf2 [M.musc	4.78	10.50
	322378	AF054819	Hs.201877	DESC1 protein	1.00	1.00
	322388	AI815730	Hs.247474	hypothetical protein FLJ21032	7.09	8.49
	322416	AA223183	Hs.298442	adaptor-related protein complex 3, mu 1	3.20	5.80
50	322419	AA248987	Hs.14084	ring finger protein 7	1.64	1.57
	322425	W37943	Hs.34892	KIAA1323 protein	0.83	1.00
	322431	AA069222	Hs.141892	ESTs	3.96	5.22
	322450	AA040131	Hs.25144	ESTs	5.18	12.67
	322465	AA137152	Hs.286049	phosphoserine aminotransferase	3.41	2.23
55	322467	AF116826	Hs.180340	putative protein-tyrosine kinase	1.00	1.30
	322473	AA744286	Hs.266935	tRNA selenocysteine associated protein	1.75	2.03
	322508	T52172	Hs.302213	ESTs	1.00	2.27
	322523	W80398	Hs.193197	ESTs	2.75	5.49
	322527	AF147359		gb:Homo sapiens full length insert cDNA	1.25	1.27
60	322560	AI916847	Hs.270947	ESTs	4.57	8.81
	322566	W87285	Hs.269587	ESTs	1.00	1.42
	322585	AA837622		gb:zh69c01.r1 Soares_fetal_liver_spleen_	4.18	6.94
	322635	AA679084		gb:zh90h08.r1 Soares_fetal_liver_spleen_	2.40	4.85
	322641	AA007352	Hs.256042	ESTs	2.94	4.64
65	322653	AI828854	Hs.258538	striatin, calmodulin-binding protein	0.48	0.38
	322664	AA011522		gb:zl03g07.r1 Soares_fetal_liver_spleen_	1.92	2.18
	322687	AI110759		gb:AF074666 Human fetal liver cDNA libra	4.14	6.75
	322692	AA018117	Hs.60843	potassium voltage-gated channel, shaker-	3.50	5.00
	322694	AI110872	Hs.279812	PRO0327 protein	1.80	1.72
70	322708	AF113674	Hs.283773	clone FLB1727	1.00	3.43
	322712	AA021328	Hs.23507	hypothetical protein FLJ11109	3.28	3.86
	322766	AW068805	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.63	1.53
	322770	AA045796	Hs.122682	ESTs	1.53	1.06
	322794	AI608591	Hs.38991	S100 calcium-binding protein A2	12.06	1.94
75	322810	AI962276	Hs.127444	ESTs	4.09	6.90
	322818	AW043782	Hs.293616	ESTs	1.20	1.63
	322820	AI377755	Hs.120695	ESTs	0.21	1.93
	322872	AA827228	Hs.126943	ESTs	2.04	1.63
	322882	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	5.26	1.22
80	322887	AI986306	Hs.86149	phosphoinositol 3-phosphate-binding prot	2.80	2.24
	322913	AI733737	Hs.68837	ESTs	2.38	6.61
	322926	AI825940	Hs.211182	ESTs	4.02	5.79
	322929	AI365585	Hs.146246	ESTs	0.30	1.14
	322968	AI905228	Hs.83484	SRY (sex determining region Y)-box 4	2.05	1.13
85	322971	C15953	Hs.212760	hypothetical protein FLJ13649	1.18	2.00
	322981	AA493252	Hs.159577	ESTs	2.28	2.61

	322988	C18727	Hs.171941	ESTs	0.39	2.00
	323003	A1733859	Hs.149089	ESTs	3.28	1.00
	323013	AA134042	Hs.191451	ESTs	3.38	5.68
5	323025	AL157565	Hs.315369	Homo sapiens cDNA: FLJ23075 fis, clone L	0.05	1.10
	323032	AW244073	Hs.145946	ESTs	10.18	21.27
	323052	R21124	Hs.85573	Homo sapiens DC29 mRNA, complete cds	1.45	1.90
	323064	AL119341	Hs.49359	Homo sapiens mRNA; cDNA DKFZp547E052 (fr	3.08	5.64
	323098	AL700025	Hs.270471	ESTs	2.31	4.49
10	323102	AL119913	Hs.163615	ESTs	5.38	11.64
	323155	AL135041		gb:DKFZp762K2310_r1 762 (synonym: hmel2)	2.38	5.56
	323176	AW071648	Hs.82101	pleckstrin homology-like domain, family	1.05	1.41
	323191	AA195600	Hs.301570	ESTs	0.73	1.24
	323225	AA205654	Hs.24790	KIAA1573 protein	5.25	11.95
15	323232	AA148722	Hs.224680	ESTs	0.45	1.35
	323266	AW003362	Hs.243885	nuclear autoantigenic sperm protein (his	1.71	1.83
	323281	A0697556	Hs.292659	ESTs	1.24	3.21
	323283	AA256014	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	12.68	15.05
	323314	AA226310	Hs.191501	ESTs	4.42	9.61
20	323316	AL134620	Hs.280175	ESTs	2.98	5.93
	323334	AL336501	Hs.77273	ras homolog gene family, member A	1.98	3.30
	323338	R74219	Hs.23348	S-phase kinase-associated protein 2 (p45	1.62	1.00
	323348	AA233056	Hs.191518	ESTs	1.00	1.07
	323351	AA704103	Hs.24049	ESTs	1.43	1.68
	323359	AA234172	Hs.137418	ESTs	0.34	1.18
25	323360	AA716081	Hs.161719	ESTs	3.01	3.71
	323405	AW139550	Hs.115173	ESTs	1.90	8.81
	323420	A1672386	Hs.263780	ESTs	0.29	1.01
	323434	AW081455	Hs.120219	ESTs	2.27	1.92
30	323445	AA253103	Hs.135569	ESTs, Weakly similar to NEUROD (H.sapien	0.43	0.80
	323449	AA282865	Hs.284153	Fanconi anemia, complementation group A	3.19	3.85
	323492	H00978	Hs.20887	hypothetical protein FLJ10392	2.70	3.20
	323501	AA182461	Hs.84520	ESTs	2.04	3.31
	323505	A1652287		gb:EST382593 MAGE resequences, MAGK Homo.21		3.08
35	323515	AA282274	Hs.256083	ESTs	2.69	3.40
	323541	A185116	Hs.104613	RP42 homolog	1.20	1.09
	323545	A1814405	Hs.224569	ESTs	1.25	1.55
	323635	R63117	Hs.9591	Homo sapiens cDNA: FLJ23249 fis, clone C	0.27	0.72
	323675	AA984759	Hs.272168	tumor differentially expressed 1	3.70	5.80
40	323678	AL042121	Hs.20880	ESTs	3.33	5.10
	323691	AA317561	Hs.145599	ESTs	1.00	1.00
	323693	AW297758	Hs.249721	ESTs	2.01	1.54
	323746	AW298611	Hs.12808	MARK	4.11	5.53
	323774	AA329806	Hs.321056	Homo sapiens mRNA; cDNA DKFZp586F1322 (f	2.06	3.70
45	323856	AA355264	Hs.267604	hypothetical protein FLJ10450	3.42	8.13
	323857	T18998	Hs.293668	ESTs	5.97	12.51
	323870	AA341774	Hs.129212	ESTs	3.17	4.52
	323876	AL042492	Hs.147313	ESTs	0.36	1.00
	323885	AA344308	Hs.128427	Homo sapiens BAC clone RP11-335J18 from	2.31	3.33
50	323911	AL043212	Hs.92550	ESTs	4.38	5.41
	323919	AA862973	Hs.220704	ESTs	5.80	10.20
	323972	A1869964	Hs.182906	ESTs	3.10	5.14
	324005	AA610011	Hs.208021	ESTs	5.34	10.07
	324036	A1472078	Hs.303662	ESTs	1.00	5.03
55	324055	AA528794	Hs.128644	ESTs	0.86	1.00
	324063	AW292740	Hs.272813	dual oxidase 1	0.45	0.91
	324072	AA381829		gb:EST94855 Activated T-cells I Homo sap	2.82	5.12
	324092	AW269931	Hs.202473	Homo sapiens cDNA: FLJ22278 fis, clone H	2.40	2.52
	324095	AW377983	Hs.298140	Homo sapiens cDNA: FLJ22502 fis, clone H	1.32	4.30
60	324129	A1381918	Hs.285833	Homo sapiens cDNA: FLJ22135 fis, clone H	1.40	1.77
	324132	AW504860	Hs.288836	hypothetical protein FLJ12673	4.24	6.21
	324214	AA412395	Hs.225740	ESTs	6.95	10.69
	324227	AA295552	Hs.28631	Homo sapiens cDNA: FLJ22141 fis, clone H	0.81	0.53
	324266	AL047634	Hs.231913	ESTs	2.42	4.05
65	324275	AA429088	Hs.98523	ESTs	3.62	5.38
	324281	AL048026	Hs.124675	ESTs, Weakly similar to T14742 hypotheti	0.14	0.70
	324280	AA432032	Hs.304420	ESTs	3.71	4.34
	324303	AL118754		gb:DKFZp761P1910_r1 761 (synonym: hamy2)	0.95	0.91
	324312	A1198841	Hs.128173	ESTs	4.06	5.91
70	324325	AL138153	Hs.300410	ESTs	5.88	8.25
	324338	AL138357	Hs.145078	regulator of differentiation (In S. pombe	0.87	1.25
	324341	AW197734	Hs.99807	ESTs, Weakly similar to unnamed protein	1.28	1.00
	324343	AW452016	Hs.293232	ESTs	2.54	3.46
	324371	AA452305	Hs.270319	ESTs	5.85	8.36
75	324382	AW502749	Hs.24724	MFH-amplified sequences with leucine-ric	0.76	1.64
	324384	AA453396	Hs.127656	KIAA1349 protein	2.88	5.69
	324385	F28212	Hs.284247	KIAA1491 protein	1.81	1.99
	324388	A1924963	Hs.306206	hypothetical protein FLJ11215	1.00	1.00
	324432	AA464510	Hs.152812	ESTs	2.73	2.17
	324497	AW152624	Hs.136340	ESTs, Weakly similar to unnamed protein	0.71	1.90
80	324510	A1148353	Hs.287425	Homo sapiens cDNA FLJ11569 fis, clone HE	1.00	1.00
	324580	AA492588		gb:ng99c08.s1 NCL CGAP_Thy1 Homo sapiens	2.18	3.50
	324582	AA506935	Hs.132036	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.96	11.36
	324633	AA572994	Hs.325489	ESTs	2.92	4.22
85	324640	AW295832	Hs.134798	ESTs, Moderately similar to TTL MOUSE TU	5.48	11.74
	324675	AW014734	Hs.157969	ESTs	0.39	0.73



	324699	AW504732	Hs.21275	hypothetical protein FLJ11011	0.93	0.93
	324747	AA603532	Hs.130807	ESTs	1.57	1.81
	324748	AA657457	Hs.292385	ESTs	1.55	1.34
5	324801	AI819924	Hs.14553	sterol O-acyltransferase (acyl-Coenzyme	1.00	6.56
	324804	AI692552		gb:wd73f12.x1 NCI_CGAP_Lu24 Homo sapiens	1.00	7.53
	324828	AA843926	Hs.124434	ESTs	2.00	3.25
	324855	AW152305	Hs.122364	ESTs	2.74	3.43
	324866	AI541214	Hs.46320	Small proline-rich protein SPRK (human,	1.07	0.95
10	324871	AW297755	Hs.271923	Homo sapiens cDNA: FLJ22785 fis, clone K	1.68	1.21
	324886	AA806794	Hs.131511	ESTs	2.56	5.61
	324889	D31010		gb:HUM1.12147 Human fetal lung Homo sapia	2.20	4.65
	324948	AW383618	Hs.265459	ESTs, Moderately similar to ALU2_HUMAN A	5.28	7.05
	324953	AI264628	Hs.125428	ESTs	3.37	6.51
15	324958	AA625076	Hs.132892	protocadherin 20	5.12	9.81
	324988	T06997	Hs.121028	hypothetical protein FLJ10549	2.52	1.08
	325024	F13254	Hs.78672	laminin, alpha 4	5.24	10.22
	325105	H97109	Hs.105421	ESTs	1.00	1.00
	325108	AA401853	Hs.22380	ESTs	1.99	2.14
20	325114	D83901	Hs.315562	ESTs	2.73	3.17
	325146	AI064690	Hs.171176	ESTs	1.86	3.41
	325149	D61117	Hs.187646	ESTs	0.42	0.93
	325187	AI653682	Hs.197812	ESTs	6.50	11.31
	325228				6.18	15.76
25	325235				2.64	4.12
	325328				2.87	4.42
	325340				0.29	0.33
	325367				16.56	24.29
	325373				0.63	1.22
30	325389				0.88	1.05
	325436				5.75	14.14
	325471				8.46	17.82
	325498				3.32	6.42
	325557				5.51	8.28
35	325559				7.48	21.40
	325560				4.08	6.25
	325569				4.20	5.24
	325585				1.10	1.13
	325587				1.00	1.00
40	325597				2.98	13.40
	325639				0.78	0.78
	325685				0.46	0.66
	325686				0.95	1.55
	325735				4.48	9.20
45	325739				0.59	0.88
	325740				2.42	6.61
	325792				7.88	9.83
	325819				4.74	7.18
	325883				2.02	2.64
50	325895				7.78	15.98
	325925				2.04	10.60
	325932				4.18	7.35
	325941				3.66	9.03
	325969				0.61	0.80
55	325971				4.88	7.42
	326025				0.55	1.07
	326046				7.21	14.72
	326099				3.60	5.98
	326108				1.27	1.06
60	326163				3.27	5.70
	326165				0.45	1.11
	326189				0.13	0.45
	326204				5.60	9.00
	326230				7.00	12.01
65	326274				1.00	8.09
	326360				9.86	15.35
	326393				0.52	0.77
	326505				1.00	1.42
	326515				1.24	5.84
70	326589				9.20	13.49
	326592				2.77	4.01
	326605				2.01	2.53
	326692				1.00	1.00
	326693				1.00	1.31
75	326720				0.19	0.65
	326742				2.34	7.20
	326770				0.25	0.83
	326818				3.09	4.56
	326936				2.08	3.45
80	326964				0.41	1.70
	326983				2.02	3.80
	326991				1.09	1.20
	327036				1.00	8.04
	327040				3.05	4.22
85	327053				3.55	6.31
	327075				1.59	1.40

WO 02/086443

PCT/US02/12476

	327085	2.50	12.57
	327130	5.38	8.04
	327155	3.74	6.58
5	327220	1.28	1.54
	327224	6.56	12.91
	327288	2.61	5.40
	327321	2.42	3.11
	327332	6.62	10.58
10	327361	2.69	4.41
	327377	2.04	6.72
	327396	2.61	4.50
	327414	1.00	8.01
	327442	5.91	9.65
15	327467	6.58	18.01
	327473	3.79	7.48
	327483	4.08	8.87
	327562	0.68	2.86
	327568	1.00	2.00
20	327606	2.06	3.61
	327611	5.90	14.26
	327642	4.06	8.74
	327654	1.05	2.08
	327734	1.00	1.00
25	327775	1.46	11.79
	327796	3.47	5.65
	327840	3.26	6.64
	327940	5.84	15.58
	327984	0.36	1.50
30	328004	1.87	1.42
	328021	0.42	0.59
	328068	2.83	4.68
	328100	3.04	5.39
	328101	3.54	5.20
35	328113	0.72	0.91
	328157	5.58	5.16
	328196	5.76	11.13
	328197	5.99	10.58
	328264	3.11	4.88
40	328299	2.20	3.06
	328342	1.49	1.94
	328365	1.00	1.00
	328369	4.40	7.36
	328381	1.86	4.93
45	328451	5.51	7.56
	328481	0.13	0.72
	328500	2.71	3.97
	328530	5.41	7.62
	328600	3.14	10.68
50	328608	4.58	8.17
	328616	2.24	11.91
	328623	3.04	5.46
	328632	0.70	1.19
	328664	3.48	6.80
55	328666	10.42	26.47
	328698	9.68	14.56
	328700	2.74	10.22
	328708	0.15	0.57
	328735	6.23	8.91
60	328743	3.62	6.54
	328806	0.22	0.78
	328861	3.68	10.54
	328908	5.42	16.36
	328933	2.02	5.29
65	328934	1.73	4.45
	328949	3.34	5.41
	329005	2.88	7.26
	329011	2.52	3.72
	329033	1.00	1.03
70	329037	5.07	8.16
	329067	1.98	2.41
	329134	2.24	3.25
	329157	2.30	11.04
	329178	2.64	5.02
75	329192	6.41	15.27
	329194	0.31	0.79
	329204	1.60	3.75
	329224	2.99	6.11
	329228	0.83	0.83
80	329288	0.63	1.01
	329337	1.00	1.00
	329541	0.76	1.68
	329560	1.34	2.02
	329588	1.68	2.22
85	329643	4.18	11.77
	329703	1.00	1.00

	329764			5.78	15.50
	329816			2.09	5.44
	329860			3.13	10.77
	329993			7.83	14.21
5	330020			5.58	13.12
	330036			3.32	5.57
	330052			4.31	7.97
	330085			1.34	1.76
	330088			4.70	12.46
10	330093			0.44	1.06
	330100			3.47	4.83
	330105			2.14	3.61
	330107			3.17	6.87
	330120			5.61	11.89
15	330123			4.50	12.74
	330208			1.55	7.62
	330263			13.10	23.38
	330300			2.81	4.98
	330313			3.00	4.41
20	330366			0.67	0.76
	330372			4.76	11.82
	330385	AA449749	Hs.182971	2.14	2.15
	330397	D14659	Hs.154387	0.40	1.15
	330468	L10343	Hs.112341	1.11	0.94
25	330472	L24203	Hs.82237	1.67	1.17
	330478	L38486	Hs.296049	0.46	1.07
	330493	M27826	Hs.267319	1.07	0.95
	330495	M31328	Hs.71642	0.97	0.96
	330506	M61906	Hs.6241	0.17	3.66
30	330512	M80563	Hs.81256	0.60	1.06
	330537	U19765	Hs.2110	2.81	2.07
	330547	U32989	Hs.183671	3.91	1.49
	330551	U39840	Hs.298867	1.15	1.03
	330568	U56244		2.83	4.79
35	330599	U90437		2.08	1.54
	330601	U90916	Hs.82845	0.89	1.35
	330605	X02419	Hs.77274	1.87	1.55
	330609	X04741	Hs.76118	1.83	1.30
	330617	X53567	Hs.85266	1.54	1.15
40	330630	X78659	Hs.79088	1.39	1.19
	330644	Y07755	Hs.38991	3.83	1.13
	330650	Z68228	Hs.2340	1.25	0.95
	330660	AA347868	Hs.139293	15.50	29.07
	330692	AA017045	Hs.6702	1.00	1.00
45	330707	AA133891	Hs.293690	0.20	1.35
	330715	AA233707	Hs.11571	0.12	1.40
	330717	AA233926	Hs.52620	6.62	5.42
	330722	AA243560	Hs.34382	1.40	1.65
50	330740	AA297746	Hs.22654	0.27	2.04
	330742	AA400979	Hs.25691	0.44	0.90
	330744	AA406142	Hs.12393	0.71	3.23
	330751	AA428286	Hs.29643	1.66	1.52
	330760	AA448663	Hs.30469	0.52	0.90
55	330763	AA450200	Hs.274337	0.37	0.97
	330786	D60374	Hs.49136	0.78	0.84
	330789	T48536	Hs.105807	0.23	3.17
	330814	AA015730	Hs.265398	0.37	2.07
	330827	AA040332	Hs.12744	1.60	1.00
60	330844	AA063037	Hs.66803	0.93	1.16
	330901	AA157818	Hs.287319	1.02	1.03
	330931	F01443	Hs.284256	0.24	0.88
	330952	H02855	Hs.29567	0.08	1.31
	330961	H10998	Hs.7164	1.29	1.26
65	330968	H16568	Hs.23748	0.48	0.96
	331014	H98597	Hs.30340	0.29	0.74
	331046	N66563	Hs.191358	0.99	8.56
	331060	N75081	Hs.157148	1.24	1.00
	331099	R36671	Hs.83937	0.75	1.03
70	331108	R41408	Hs.21983	1.00	2.75
	331131	R54797		6.04	10.68
	331135	R61398	Hs.4197	0.80	0.96
	331170	T23461	Hs.159293	2.63	4.29
	331180	T32446	Hs.6640	1.78	2.71
	331183	T40769	Hs.8469	1.00	3.01
75	331203	T82310		1.70	3.80
	331271	AA059347	Hs.82226	1.20	3.19
	331306	AA252079	Hs.63931	0.31	1.30
	331327	AA281076	Hs.109221	2.09	2.41
	331341	AA303125	Hs.23240	0.72	2.43
80	331359	AA416979	Hs.46901	0.09	0.91
	331363	AA421562	Hs.91011	1.02	0.87
	331378	AA448881	Hs.49282	1.03	1.23
	331384	AA456001	Hs.93847	1.40	1.00
	331402	AA505135	Hs.44037	1.80	3.93
85	331422	F10802	Hs.163628	1.65	1.89

	331490	N32912	Hs.26813	CDA14	2.48	1.73
	331531	N51343		gbxyz15g04.s1 Soares_multiple_sclerosis_	0.98	1.68
	331547	N54811		gbxod74f04.s1 NCI_CGAP_Ov2 Homo sapiens	3.80	5.75
5	331578	N67960	Hs.249989	ESTs	0.11	0.67
	331589	N71027	Hs.152618	ESTs	1.09	1.38
	331608	N89861	Hs.112110	PTD007 protein	0.93	0.76
	331614	N92293	Hs.240272	EST	0.17	1.34
	331668	W69707	Hs.58030	EST	2.24	3.82
10	331671	W72033	Hs.194695	ras homolog gene family, member I	1.00	1.24
	331676	W79834	Hs.58559	ESTs, Weakly similar to rholekin [M.musc	0.08	1.07
	331681	W85712	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	8.72	4.27
	331692	W93592	Hs.152213	wingless-type MMTV Integration site faml	0.94	0.54
	331717	AA190888	Hs.153881	Homo sapiens NY-REN-62 antigen mRNA, par	1.57	1.34
15	331718	AA191404	Hs.104072	ESTs	6.80	11.77
	331811	AA404500	Hs.301570	ESTs	1.10	1.00
	331820	AA405970	Hs.97996	transcription termination factor, mitoc	0.73	0.59
	331831	AA412031	Hs.97901	EST	2.77	4.08
	331852	AA418988	Hs.98314	Homo sapiens mRNA; cDNA DKFZp586L0120 (f	0.23	0.93
20	331943	AA453418	Hs.21275	hypothetical protein FLJ11011	0.36	1.88
	331969	AA450702	Hs.82772	collagen, type XI, alpha 1	1.00	1.00
	331990	AA478102	Hs.139631	ESTs	3.04	3.87
	332002	AA482009	Hs.105104	ESTs	1.19	0.78
	332027	AA489671	Hs.65641	hypothetical protein FLJ20073	1.27	1.03
25	332029	AA489697	Hs.145053	ESTs	0.30	1.62
	332033	AA489840	Hs.251014	EST	2.30	3.70
	332048	AA495019	Hs.201591	ESTs	0.17	0.52
	332071	AA598594	Hs.205293	KIAA1211 protein	1.35	1.23
	332074	AA599012		gbxae41e11.s1 Gessler Wilms tumor Homo s	0.19	2.00
30	332083	AA600200	Hs.155546	KIAA1080 protein; Golgi-associated, gamm	0.31	1.18
	332085	AA600353	Hs.173933	nuclear factor I/A	0.30	1.50
	332125	AA609861	Hs.312447	ESTs	0.22	0.62
	332177	F10812	Hs.101433	ESTs	8.21	18.03
	332180	H03348	Hs.7327	claudin 1	2.27	1.57
35	332185	H10356	Hs.101689	ESTs	0.09	1.18
	332203	H49388	Hs.317769	EST	8.05	5.02
	332232	N48891	Hs.101915	Stargardt disease 3 (autosomal dominant)	0.78	0.85
	332240	N54803	Hs.324267	ESTs, Weakly similar to putative p150 [	0.96	1.23
	332261	N70294	Hs.269137	ESTs	2.40	3.74
40	332275	R08838	Hs.26530	serum deprivation response (phosphatidyl	0.27	0.75
	332280	R38100	Hs.146381	RNA binding motif protein, X chromosome	0.39	1.88
	332299	R69250	Hs.21201	necln 3; DKFZP566B0846 protein	5.24	12.76
	332304	R74041	Hs.101539	ESTs	1.44	3.18
	332314	T25862	Hs.101774	hypothetical protein FLJ23045	0.68	1.32
45	332384	M11433	Hs.101850	retinol-binding protein 1, cellular	1.71	0.88
	332434	N75542	Hs.289068	Homo sapiens cDNA FLJ11918 fis, clone HE	0.43	0.86
	332445	T63781	Hs.11112	ESTs	0.68	1.00
	332453	L00205	Hs.111758	keratin 6A	31.54	1.00
	332458	M33493	Hs.250700	trypsin beta 1	0.51	1.00
50	332504	AA053917	Hs.15106	chromosome 14 open reading frame 1	0.79	1.24
	332525	M17252	Hs.278430	cytochrome P450, subfamily XXIA (steroid	0.98	1.70
	332530	M31682	Hs.1735	inhibin, beta B (activin AB beta polypep	0.88	0.66
	332535	N20284	Hs.19280	cysteine-rich motor neuron 1	0.22	1.46
	332539	AA412528	Hs.20183	ESTs, Weakly similar to AF164793 1 prote	0.93	1.49
55	332559	M13955	Hs.166189	cytokeratin 2	0.35	1.13
	332563	N92924	Hs.274407	protease, serine, 16 (thymus)	1.00	1.00
	332565	AA234896	Hs.25272	E1A binding protein p300	0.36	1.05
	332594	AA279313	Hs.3239	methyl CpG binding protein 2 (Rett syndr	0.53	0.59
	332634	S38953	Hs.283750	tenascin XA	0.38	1.16
60	332638	AA283034	Hs.50640	JAK binding protein	1.00	1.70
	332640	AA417152	Hs.5101	protein regulator of cytokinesis 1	6.15	1.16
	332654	AA001296	Hs.288217	hypothetical protein MGC2941	1.50	2.73
	332665	AA223335	Hs.63788	propionyl Coenzyme A carboxylase, beta p	1.20	0.91
	332692	AA496035	Hs.247926	gap junction protein, alpha 5, 40kD (con	0.17	1.12
65	332716	L00058	Hs.79070	v-myc avian myelocytomatosis viral onco	1.00	1.44
	332736	L13773	Hs.114765	myeloid/lymphoid or mixed-lineage leukem	1.00	1.81
	332758	X93921	Hs.296938	dual specificity phosphatase 7	0.53	0.78
	332781	AA233258	Hs.247112	hypothetical protein FLJ10902	1.44	1.56
	332792				1.70	1.19
70	332816				1.85	2.47
	332858				1.04	1.57
	332906				3.48	8.04
	332911				1.00	1.00
	332912				1.06	4.40
75	332922				1.00	1.00
	332956				0.42	0.88
	332959				1.96	6.34
	332982				0.56	0.89
	332984				0.30	0.78
80	332998				1.47	2.01
	333058				0.47	1.38
	333097				2.14	3.19
	333121				2.76	3.70
	333122				1.92	1.21
85	333123				1.85	1.39
	333138				0.47	0.52

	333139	1.88	0.84
	333140	0.21	0.64
	333221	1.51	1.11
	333260	0.75	1.01
5	333380	6.68	15.75
	333387	4.56	12.61
	333512	5.05	8.01
	333524	2.28	3.98
10	333585	2.31	1.53
	333603	2.23	1.17
	333604	2.51	1.58
	333618	0.52	0.98
	333627	1.44	1.35
	333628	1.90	1.90
15	333650	1.85	2.10
	333678	1.85	2.35
	333750	2.18	5.67
	333763	1.99	2.60
	333767	1.02	0.96
20	333768	1.78	1.65
	333769	2.15	2.13
	333772	1.46	2.53
	333777	1.00	1.42
	333846	2.99	4.50
25	333884	0.47	0.94
	333887	0.50	1.00
	333891	0.43	0.89
	333892	0.51	0.91
	333904	0.26	1.13
30	333906	0.55	0.98
	333948	1.70	2.15
	333954	0.37	1.09
	333966	8.10	14.30
	333968	0.63	1.38
35	334051	4.24	12.30
	334094	1.30	12.03
	334113	4.55	8.63
	334161	0.82	1.59
	334183	0.47	0.76
40	334187	1.36	3.70
	334219	0.69	1.04
	334222	1.88	1.70
	334223	4.72	3.14
	334239	0.79	0.62
45	334255	0.45	1.10
	334333	1.00	3.56
	334378	3.98	5.76
	334382	1.50	1.31
	334492	3.59	4.75
50	334562	5.94	15.40
	334588	8.14	19.53
	334616	1.55	1.55
	334633	5.16	8.07
	334648	0.59	2.13
55	334787	3.70	7.15
	334866	8.13	10.60
	334891	0.32	1.14
	334933	1.00	3.84
	334934	4.01	7.43
60	334945	1.04	2.96
	334967	0.29	1.14
	334990	1.50	1.39
	335015	5.88	18.65
	335093	0.55	1.75
65	335120	4.31	8.01
	335125	0.38	1.97
	335179	1.24	1.99
	335188	0.46	1.47
	335211	1.61	1.42
70	335288	0.73	0.97
	335289	0.20	0.26
	335361	2.18	1.58
	335379	0.50	0.71
	335414	3.64	14.94
75	335416	2.93	3.98
	335498	0.95	0.91
	335497	1.71	1.92
	335548	1.15	2.40
	335551	3.22	10.54
80	335558	3.42	4.89
	335586	5.50	12.75
	335619	2.99	3.07
	335620	3.80	8.29
85	335621	0.28	0.57
	335682	0.46	1.17

WO 02/086443

PCT/US02/12476

	335686	2.55	3.81
	335755	2.24	1.07
	335784	0.20	0.97
5	335814	1.13	1.48
	335815	2.45	3.51
	335823	1.00	4.16
	335835	0.49	1.70
	335851	1.66	1.39
10	335868	2.98	6.43
	335896	0.98	0.99
	335936	12.10	21.93
	335948	1.00	1.54
	335983	1.00	4.21
15	335995	0.37	1.17
	336021	1.04	0.84
	336034	11.40	23.54
	336038	1.19	1.21
	336066	0.54	1.63
20	336107	0.95	0.70
	336205	3.13	6.29
	336275	3.20	10.10
	336292	2.34	3.09
	336331	1.00	1.00
25	336419	0.65	0.79
	336632	2.33	2.16
	336633	2.55	2.23
	336634	2.19	2.03
	336635	2.69	2.48
30	336636	2.13	1.83
	336637	2.43	2.24
	336638	2.31	2.03
	336659	0.60	1.31
	336675	0.31	1.18
35	336684	1.50	1.14
	336694	4.74	7.10
	336716	4.43	6.37
	336721	2.20	0.74
	336798	1.64	2.14
40	336900	6.14	12.73
	336948	1.00	1.00
	337028	1.30	2.09
	337043	4.01	11.53
	337046	1.67	1.84
45	337054	2.78	7.35
	337128	7.20	16.14
	337162	3.45	5.34
	337183	5.72	11.41
	337184	3.72	5.90
50	337192	1.27	1.06
	337194	1.88	1.68
	337229	0.22	1.03
	337268	1.00	3.31
	337299	3.23	5.14
55	337325	2.76	3.72
	337389	5.80	10.42
	337493	2.06	6.30
	337497	7.88	20.29
	337500	3.80	4.48
60	337549	1.66	2.31
	337603	1.27	8.54
	337605	5.76	7.16
	337671	0.73	0.97
	337755	1.54	0.92
65	337786	5.07	9.73
	337809	6.18	12.87
	337862	3.78	12.97
	337871	2.66	8.16
	337958	0.26	1.34
70	338008	1.48	1.12
	338033	2.38	14.59
	338083	0.65	2.16
	338110	1.00	1.61
	338112	5.86	8.25
75	338145	1.70	1.97
	338148	8.07	18.19
	338158	1.30	4.55
	338161	2.58	3.57
	338179	1.00	1.00
80	338182	3.32	4.63
	338189	1.00	3.34
	338197	0.99	1.69
	338199	4.58	7.62
	338215	6.01	15.85
85	338279	0.53	0.95
	338316	20.58	38.66

	338322	3.23	7.39
	338357	4.10	11.39
	338359	10.12	21.59
5	338366	0.69	1.02
	338374	0.40	1.18
	338414	0.47	1.06
	338418	6.12	13.86
	338469	3.09	5.11
10	338501	6.28	10.32
	338506	6.97	12.41
	338523	3.10	5.84
	338549	1.70	2.70
	338561	0.79	0.81
15	338662	1.72	1.46
	338671	0.17	0.91
	338676	2.10	15.86
	338726	1.20	1.09
	338779	0.12	0.57
20	338804	0.99	1.67
	338836	1.00	1.00
	338871	4.30	9.81
	338872	5.02	12.81
	338879	0.23	1.12
25	338937	6.55	12.26
	338966	1.76	5.42
	338993	1.00	2.40
	339047	5.28	10.81
	339100	5.10	6.88
30	339114	1.00	1.70
	339121	1.00	3.75
	339170	10.36	19.67
	339229	4.08	13.48
	339264	2.64	3.83
35	339293	1.73	1.94

TABLE 8B shows the accession numbers for those Plays in Table 8A lacking unigenes. For each probe set we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Play: Unique Eos probe set identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Play	CAT number	Accessions
322044	187363_1	AW340926 AA249063 N86075
322060	44320_1	AI341937 AW003063 U34725 AA904742
321430	42705_1	X57414 X57415
321467	43034_1	X13075 X13076
322125	46779_1	R93901 AF075073 R93902
322166	46861_1	H69434 AF085958 H69846
322173	46873_1	H52567 H52567 AF085970 H52164
322178	46882_1	H56535 AF085980 H56712
322179	46885_1	H92891 AF085982 H92777
321577	1615102_1	H84849 H84252 H84260 H86664 H85320
321587	1615333_1	H95531 H95521 H84529
313723	111953_1	AA070412 AA102346 AA081885
320997	627492_1	H22544 H46842 AI204929
322278	47271_1	W69304 AF086283 W69200
321687	218439_1	AA625149 AA313030 AA313052 H97463
313883	129439_1	AA655089 AA135130 AA484059 AA102419 AW877765
322320	47422_1	W79150 AF086419
322339	814584_1	AI668646 AI734214 W17348
314648	293660_1	AW979268 AA878419 AA431342 AA431628
300201	682222_1	AI308300 AI308296
308897	25196_2	AI093967
323155	979809_1	AL120701 AL135041 AL121524
322527	38927_1	AF147359 T58511 T58560
322585	473768_2	W88919 W89125
300362	1574395_1	Z42308 H23514
322635	82296_1	AA005129 AA679084 AA694399
322664	85042_1	AA011522 AA702841 AA011691 AA330797
315454	380580_1	AI239464 AI239473 AA625812 AI208703
322687	37372_1	AF074666 AI110759 AF090902
314852	327472_1	AI903735 AA491283 AI694953 AW976803 AA761362
307783	697809_1	AI347274 AW844024
324072	269032_1	AA381722 AA381829 AW963906 AW963902 AA381242
300627	221345_1	AA488472 W27363 AA317053 BE082689 AW967036 BE079872
323505	196389_1	AW970512 AA280251 AI652287 BE466438 AI650725 AA551854 AA281574 AW571481
315791	403558_1	AA678177 AA677034
324303	233842_1	AL118754 AA333202 H38001
316519	442885_1	AA847835 AA768376
300926	333127_1	AA504860 AA504911

	324580	328264_1	AA492588 AA492488 AA492571
	301882	275087_1	T78054 T79888 AA398185
	324804	398093_1	AI692552 AI393343 AI800510 AI377711 F24263 AA661876
5	324889	1515978_1	D31010 D30991 D31168 D31166 D31465
	302697	43219_1	AJ001409 AJ001410
	302711	45419_1	L08442 D51348
	302742	458_39	L12081
	318499	364430_1	T25451 AA585296 AA585305
10	310624	34624_4	U88896 U88898 AA916056 T03285 AI341594 AI359534 AI634031 U88897
	302847	458_105	X98941 X98942 X98943 X98953 X98949
	304122	77271_5	H28966
	303598	270283_1	AA382814 AA402411 AA412355
	311409	837264_1	AI698839 AI909260 AI909259
15	312094	797889_1	Z78390 T97427
	319312	1540116_1	Z45481 F12393 T74437
	319407	1688823_1	R05329 R01555 R08276
	319425	1689571_1	T82930 R02424 T85145
	320007	229683_1	AA336314 T82938 AA327744 AW967388 AA639967 T10753
20	320018	1815987_1	T83263 T85731 T85730
	319484	1691553_1	T91772 R07257 R07098
	318865	1535937_1	H10818 F07831 Z43072
	312220	1671607_1	N74613 T98756 T98589
	319546	243305_1	R09692 R09414 AA346353
	312389	902067_1	AI863140 W80703 R43474
25	319611	1566863_1	H14957 R56522 R11908
	312437	291472_1	BE080180 AW827313 AW231970 AA995028 AA428584 AW872716 AW892508 AW854593 AA578441 AW975234 AA664937 AA984131 AA528743 AA552874 AA564758 AW063245 AI267534 AW070190 AW893483 AA770330 AA906928 AA906582 AA758746 AA551717 AW063311 AA429538
30	311896	579192_1	AW206447 AI248530 AI084433 AI400976 R16553
	319834	112523_1	AA071267 T65940 T64515 AA071334
	321102	80531_1	AA018306 H38925 AA001221
	321158	410938_1	H79570 H47798 AA700289
	321199	212379_1	N34524 AA305071 AW954803 AA502335 AI433430 AI203597 AW026670 AW265323 AW850787 AA317554 AW993643 AW835572 AW385512 AI334956 W32951 H62656 H53902 R88904 AW835732
35	305528	28832_3	AA769156
	321270	1662057_1	N59537 N78278 R83560
	314126	177666_1	AA226431 AA226669 AA488748
	320714	743644_1	R91883 AI445591
40	306442	AA976899	
	306446	AA977348	
	306458	AA978186	
	306510	AA988546	
	306557	AA994530	
45	306572	AA995686	
	306582	AA996248	
	306656	AI004024	
	306686	AI015615	
	306751	AI032589	
50	308011	AI439473	
	306892	AI092465	
	308106	AI476803	
	308154	AI500600	
	308956	AI125111	
55	306958	AI125152	
	308213	AI557041	
	308216	AI557135	
	308219	AI557246	
	308588	AI718299	
	308599	AI719893	
60	308643	AI745040	
	308673	AI780864	
	308697	AI767143	
	308778	AI811109	
65	308808	AI818289	
	308875	AI832332	
	308886	AI833240	
	308898	AI858845	
	308966	AI870704	
	308979	AI873111	
70	303011	41689_1	AF090405 AF090407 AF090406
	303077	44060_1	AF163305 AF163307 AF163303
	305016	AA626876	
	305034	AA630128	
75	305072	AA641012	
	305148	AA654070	
	305190	AA665955	
	303978	AW513315	
	303990	AW515465	
80	303998	AW516449	
	303999	AW516611	
	305235	AA670480	
	305312	AA700201	
	305413	AA724659	
	305447	AA737856	
85	321244	29327_1	AF068654 AF068656 AF068655



	305614	AA782866	
	305637	AA805124	
	305639	AA806138	
5	305650	AA807709	
	305690	AA813477	
	305728	AA828209	
	305759	AA835353	
	305792	AA845256	
10	307041	AI144243	
	307091	AI167439	
	307181	AI189251	
	305901	AA872968	
	305910	AA875981	
15	307415	AI242118	
	307426	AI243364	
	307517	AI275055	
	307551	AI281556	
	307581	AI282207	
20	307608	AI290295	
	307691	AI318285	
	307730	AI336092	
	307760	AI342387	
	307764	AI342731	
25	307796	AI350556	
	309045	AI910902	
	309051	AI911975	
	307807	AI351789	
	307808	AI351826	
30	307820	AI355761	
	307852	AI365541	
	309122	AI928178	
	309164	AI937761	
	309177	AI951118	
35	307902	AI380462	
	309299	AW003478	
	309303	AW004823	
	309476	AW129368	
	309532	AW151119	
40	309747	AW264889	
	309769	AW272346	
	309799	AW276964	
	309866	AW299916	
	302679	311853_1	H65022 AA186889
45	309923	AW340684	
	309928	AW341418	
	309931	AW341683	
	309933	AW341936	
	302705	31765_1	U09060 U09061
50	302789	34161_1	AJ245067 AJ245070
	304006	AW517947	
	304024	T03036	
	304026	T03160	
	304028	T03266	
55	304046	T54803	
	304061	T61521	
	304063	T62536	
	302802	34487_1	Y08250 Y08245
	304114	R78946	
60	304155	H58696	
	304203	N56929	
	304234	W81608	
	304348	AA179868	
	304430	AA347682	
65	304456	AA411240	
	304521	AA464716	
	304526	AA476427	
	304607	AA513322	
	304735	AA576453	
70	304760	AA580401	
	306015	AA897116	
	306063	AA906316	
	306065	AA906725	
	306104	AA910956	
75	306109	AA911861	
	306242	AA932805	
	306288	AA936900	
	306396	AA970223	
	330568	NOT_FOUND_entrez	U56244
80	330599	15323_-12	U90437
	331131	genbank_R54797	R54797
	331203	NOT_FOUND_entrez	T82310
	331531	genbank_N51343	N51343
	331547	467398_1	AA828597 N54811
85	332074	genbank_AA599012	AA599012

TABLE 8C shows the genomic position for those Pkeys in Table 8A lacking unigene ID's and accession numbers. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

5	Pkey:	Unique number corresponding to an Eos probaset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
10	Nt_position:	Indicates nucleotide positions of predicted exons.		
	Pkey	Ref	Strand	Nt_position
	332792	Dunham, I. et al.	Plus	73381-73768
	332816	Dunham, I. et al.	Plus	358844-360030
15	332906	Dunham, I. et al.	Plus	1923101-1923205
	332911	Dunham, I. et al.	Plus	1961767-1961858
	332912	Dunham, I. et al.	Plus	1962120-1962246
	332922	Dunham, I. et al.	Plus	2009620-2009738
	332956	Dunham, I. et al.	Plus	2510528-2510658
20	332959	Dunham, I. et al.	Plus	2518145-2518213
	333138	Dunham, I. et al.	Plus	3369205-3369323
	333139	Dunham, I. et al.	Plus	3369495-3369571
	333221	Dunham, I. et al.	Plus	3978070-3978187
	333380	Dunham, I. et al.	Plus	4904775-4904846
25	333387	Dunham, I. et al.	Plus	4910935-4910997
	333512	Dunham, I. et al.	Plus	5560510-5560564
	333524	Dunham, I. et al.	Plus	5512620-5512780
	333585	Dunham, I. et al.	Plus	6234778-6234894
	333618	Dunham, I. et al.	Plus	6562391-6562566
30	333627	Dunham, I. et al.	Plus	6620584-6620903
	333628	Dunham, I. et al.	Plus	6629004-6629233
	333650	Dunham, I. et al.	Plus	6796852-6797128
	333678	Dunham, I. et al.	Plus	7068223-7068288
	333750	Dunham, I. et al.	Plus	7608165-7608234
35	333763	Dunham, I. et al.	Plus	7692491-7692630
	333767	Dunham, I. et al.	Plus	7694407-7694623
	333768	Dunham, I. et al.	Plus	7695440-7695697
	333769	Dunham, I. et al.	Plus	7696625-7696707
	333772	Dunham, I. et al.	Plus	7706773-7706902
40	333777	Dunham, I. et al.	Plus	7746805-7746916
	333846	Dunham, I. et al.	Plus	8008623-8008757
	333884	Dunham, I. et al.	Plus	8153960-8154161
	333887	Dunham, I. et al.	Plus	8154882-8155025
	333891	Dunham, I. et al.	Plus	8156437-8156709
45	333892	Dunham, I. et al.	Plus	8156825-8157001
	333948	Dunham, I. et al.	Plus	8583497-8583627
	333954	Dunham, I. et al.	Plus	8563186-8563335
	333966	Dunham, I. et al.	Plus	8655543-8655826
	333968	Dunham, I. et al.	Plus	8681004-8681241
50	334081	Dunham, I. et al.	Plus	9686941-9687077
	334094	Dunham, I. et al.	Plus	9888953-9889105
	334113	Dunham, I. et al.	Plus	10282459-10282597
	334161	Dunham, I. et al.	Plus	10599033-10599180
	334219	Dunham, I. et al.	Plus	12716160-12716384
55	334239	Dunham, I. et al.	Plus	13056569-13056693
	334333	Dunham, I. et al.	Plus	13603544-13603657
	334378	Dunham, I. et al.	Plus	13907239-13907370
	334382	Dunham, I. et al.	Plus	13915866-13916036
	334562	Dunham, I. et al.	Plus	14987847-14987940
60	334588	Dunham, I. et al.	Plus	15032740-15032817
	334616	Dunham, I. et al.	Plus	15176123-15176470
	334633	Dunham, I. et al.	Plus	15333206-15333305
	334866	Dunham, I. et al.	Plus	18872214-18872317
	334891	Dunham, I. et al.	Plus	19299770-19299944
65	334934	Dunham, I. et al.	Plus	20103970-20104058
	335015	Dunham, I. et al.	Plus	20682792-20682945
	335120	Dunham, I. et al.	Plus	21436285-21436384
	335125	Dunham, I. et al.	Plus	21441390-21441471
	335179	Dunham, I. et al.	Plus	21634405-21634526
70	335188	Dunham, I. et al.	Plus	21669118-21669328
	335211	Dunham, I. et al.	Plus	21774611-21774680
	335361	Dunham, I. et al.	Plus	22807292-22807445
	335379	Dunham, I. et al.	Plus	22899306-22899420
	335414	Dunham, I. et al.	Plus	23235546-23235684
75	335416	Dunham, I. et al.	Plus	23237354-23237465
	335496	Dunham, I. et al.	Plus	24164386-24164545
	335497	Dunham, I. et al.	Plus	24167666-24167869
	335558	Dunham, I. et al.	Plus	24740167-24740347
	335586	Dunham, I. et al.	Plus	24990333-24990497
80	335686	Dunham, I. et al.	Plus	25439839-25439920
	335784	Dunham, I. et al.	Plus	25942710-25942792
	335823	Dunham, I. et al.	Plus	26365925-26366004
	335983	Dunham, I. et al.	Plus	27938968-27939070
	335995	Dunham, I. et al.	Plus	28009044-28009184
85	336021	Dunham, I. et al.	Plus	28686482-28686559

	336034	Dunham, I. et al.	Plus	29014404-29014590
	336038	Dunham, I. et al.	Plus	29022963-29023165
	336107	Dunham, I. et al.	Plus	29987731-29987869
5	336632	Dunham, I. et al.	Plus	983890-985529
	336633	Dunham, I. et al.	Plus	985591-986221
	336634	Dunham, I. et al.	Plus	986296-986670
	336635	Dunham, I. et al.	Plus	987908-988364
	336636	Dunham, I. et al.	Plus	988418-989185
10	336637	Dunham, I. et al.	Plus	989276-990813
	336638	Dunham, I. et al.	Plus	991906-993240
	336659	Dunham, I. et al.	Plus	1896402-1896478
	336694	Dunham, I. et al.	Plus	2420546-2420616
	336721	Dunham, I. et al.	Plus	3371522-3371586
15	336900	Dunham, I. et al.	Plus	10236423-10236523
	336948	Dunham, I. et al.	Plus	12692290-12692381
	337028	Dunham, I. et al.	Plus	16644817-16644942
	337054	Dunham, I. et al.	Plus	17821742-17821922
	337162	Dunham, I. et al.	Plus	23478943-23479145
20	337183	Dunham, I. et al.	Plus	23943606-23943696
	337184	Dunham, I. et al.	Plus	23973949-23974016
	337268	Dunham, I. et al.	Plus	28011979-28012034
	337299	Dunham, I. et al.	Plus	29022656-29022775
	337389	Dunham, I. et al.	Plus	31401509-31401579
25	337493	Dunham, I. et al.	Plus	33330760-33330981
	337549	Dunham, I. et al.	Plus	34474472-34474531
	337755	Dunham, I. et al.	Plus	3971764-3971900
	337809	Dunham, I. et al.	Plus	4449069-4449193
30	337871	Dunham, I. et al.	Plus	5443027-5443101
	337958	Dunham, I. et al.	Plus	6969162-6969270
	338008	Dunham, I. et al.	Plus	7697068-7697236
	338033	Dunham, I. et al.	Plus	8092128-8092271
	338110	Dunham, I. et al.	Plus	10384481-10384621
	338112	Dunham, I. et al.	Plus	10391398-10391600
35	338145	Dunham, I. et al.	Plus	11386629-11386692
	338148	Dunham, I. et al.	Plus	11448985-11449085
	338179	Dunham, I. et al.	Plus	12808775-12808833
	338197	Dunham, I. et al.	Plus	13638107-13638181
	338279	Dunham, I. et al.	Plus	16168944-16169091
40	338316	Dunham, I. et al.	Plus	17089711-17089888
	338322	Dunham, I. et al.	Plus	17132477-17132547
	338357	Dunham, I. et al.	Plus	18062184-18062402
	338359	Dunham, I. et al.	Plus	18074402-18074501
	338366	Dunham, I. et al.	Plus	18252026-18252189
45	338374	Dunham, I. et al.	Plus	18371200-18371282
	338414	Dunham, I. et al.	Plus	19345573-19345660
	338418	Dunham, I. et al.	Plus	19435506-19435596
	338501	Dunham, I. et al.	Plus	21244713-21244828
	338508	Dunham, I. et al.	Plus	21221871-21221953
50	338523	Dunham, I. et al.	Plus	21509763-21509864
	338662	Dunham, I. et al.	Plus	24404720-24404899
	338804	Dunham, I. et al.	Plus	27236005-27236108
	338836	Dunham, I. et al.	Plus	27792166-27792272
	338879	Dunham, I. et al.	Plus	28410653-28410734
55	338937	Dunham, I. et al.	Plus	29160655-29160725
	338993	Dunham, I. et al.	Plus	30077787-30078184
	339047	Dunham, I. et al.	Plus	30760793-30760968
	339100	Dunham, I. et al.	Plus	31141580-31141765
	339114	Dunham, I. et al.	Plus	31456454-31456519
60	339121	Dunham, I. et al.	Plus	31583467-31583536
	339170	Dunham, I. et al.	Plus	32216399-32216527
	339293	Dunham, I. et al.	Plus	33223671-33223819
	332658	Dunham, I. et al.	Minus	1339607-1339397
65	332982	Dunham, I. et al.	Minus	2628296-2628109
	332984	Dunham, I. et al.	Minus	2632606-2632457
	332998	Dunham, I. et al.	Minus	2711704-2711565
	333058	Dunham, I. et al.	Minus	3028925-3028811
	333097	Dunham, I. et al.	Minus	3204124-3204038
70	333121	Dunham, I. et al.	Minus	3308446-3308358
	333122	Dunham, I. et al.	Minus	3309596-3309531
	333123	Dunham, I. et al.	Minus	3310817-3310749
	333140	Dunham, I. et al.	Minus	3377220-3376309
	333260	Dunham, I. et al.	Minus	4308400-4308304
	333603	Dunham, I. et al.	Minus	6466335-6465727
75	333604	Dunham, I. et al.	Minus	6467090-6466768
	333904	Dunham, I. et al.	Minus	8217374-8217261
	333908	Dunham, I. et al.	Minus	8218238-8218063
	334183	Dunham, I. et al.	Minus	11832582-11832508
	334187	Dunham, I. et al.	Minus	11921456-11921205
80	334222	Dunham, I. et al.	Minus	12732417-12732289
	334223	Dunham, I. et al.	Minus	12734365-12734269
	334255	Dunham, I. et al.	Minus	13200776-13200692
	334492	Dunham, I. et al.	Minus	14478333-14478172
	334648	Dunham, I. et al.	Minus	15363301-15363222
85	334787	Dunham, I. et al.	Minus	16299093-16298937
	334933	Dunham, I. et al.	Minus	20078117-20077991

WO 02/086443

PCT/US02/12476

	334945	Dunham, I. et al.	Minus	20138885-20138637
	334967	Dunham, I. et al.	Minus	20173311-20173218
	334990	Dunham, I. et al.	Minus	20341159-20341087
5	335093	Dunham, I. et al.	Minus	21297357-21297214
	335288	Dunham, I. et al.	Minus	22304275-22303770
	335289	Dunham, I. et al.	Minus	22305550-22305708
	335548	Dunham, I. et al.	Minus	24662773-24662673
	335551	Dunham, I. et al.	Minus	24679828-24678961
10	335619	Dunham, I. et al.	Minus	25082677-25082498
	335620	Dunham, I. et al.	Minus	25092561-25092434
	335621	Dunham, I. et al.	Minus	25098878-25098767
	335682	Dunham, I. et al.	Minus	25421215-25421093
	335755	Dunham, I. et al.	Minus	25763806-25763747
15	335814	Dunham, I. et al.	Minus	26320043-26319845
	335815	Dunham, I. et al.	Minus	26320518-26320421
	335835	Dunham, I. et al.	Minus	26393311-26393245
	335851	Dunham, I. et al.	Minus	26604853-26604742
	335868	Dunham, I. et al.	Minus	26711437-26711300
20	335896	Dunham, I. et al.	Minus	26977639-26977558
	335936	Dunham, I. et al.	Minus	27360474-27360400
	335948	Dunham, I. et al.	Minus	27555924-27555788
	336066	Dunham, I. et al.	Minus	29241080-29240842
	336205	Dunham, I. et al.	Minus	30477456-30477311
25	336275	Dunham, I. et al.	Minus	32086675-32086536
	336292	Dunham, I. et al.	Minus	32818035-32817927
	336331	Dunham, I. et al.	Minus	33594527-33594371
	336419	Dunham, I. et al.	Minus	34052568-34052445
	336675	Dunham, I. et al.	Minus	2020758-2020664
30	336684	Dunham, I. et al.	Minus	2158060-2157993
	336716	Dunham, I. et al.	Minus	3259952-3259862
	336798	Dunham, I. et al.	Minus	5888954-5888757
	337043	Dunham, I. et al.	Minus	17407330-17407251
	337046	Dunham, I. et al.	Minus	17610892-17610821
35	337128	Dunham, I. et al.	Minus	22215251-22215034
	337192	Dunham, I. et al.	Minus	24591853-24591771
	337194	Dunham, I. et al.	Minus	24610510-24610359
	337229	Dunham, I. et al.	Minus	26716579-26716481
	337325	Dunham, I. et al.	Minus	30015948-30015800
40	337497	Dunham, I. et al.	Minus	33371317-33371258
	337500	Dunham, I. et al.	Minus	33376212-33376158
	337603	Dunham, I. et al.	Minus	1299296-1299194
	337605	Dunham, I. et al.	Minus	1346555-1346397
	337671	Dunham, I. et al.	Minus	3260634-3260547
45	337786	Dunham, I. et al.	Minus	4133203-4133081
	337862	Dunham, I. et al.	Minus	5347658-5347550
	338083	Dunham, I. et al.	Minus	9318438-9318301
	338158	Dunham, I. et al.	Minus	11794465-11794343
	338161	Dunham, I. et al.	Minus	12124716-12124658
50	338182	Dunham, I. et al.	Minus	12824919-12824827
	338189	Dunham, I. et al.	Minus	12878594-12878478
	338199	Dunham, I. et al.	Minus	13760885-13760780
	338215	Dunham, I. et al.	Minus	14055447-14055355
	338469	Dunham, I. et al.	Minus	20520387-20520242
55	338549	Dunham, I. et al.	Minus	22049171-22049081
	338561	Dunham, I. et al.	Minus	22311956-22311856
	338671	Dunham, I. et al.	Minus	24508421-24508346
	338676	Dunham, I. et al.	Minus	24637427-24637369
	338726	Dunham, I. et al.	Minus	25926206-25925618
60	338779	Dunham, I. et al.	Minus	27030151-27029795
	338871	Dunham, I. et al.	Minus	28301708-28301611
	338872	Dunham, I. et al.	Minus	28300921-28300790
	338966	Dunham, I. et al.	Minus	29614876-29614749
	339229	Dunham, I. et al.	Minus	32722330-32722199
65	339264	Dunham, I. et al.	Minus	32975145-32975053
	325228	6381940	Plus	2630-2694
	325235	6381943	Minus	162154-162264
	329588	3962484	Plus	1169-1619
	329560	3962491	Plus	2095-2390
70	329541	3983503	Minus	2765-3059
	325328	5866875	Plus	86780-86854
	325340	6017033	Minus	166656-166819
	325373	5866920	Minus	1136686-1136777
	325367	5866920	Minus	922881-922958
75	325389	5866921	Plus	239672-239759
	325436	5866939	Minus	29778-29907
	325498	5866967	Plus	173372-173930
	325471	6017034	Minus	289268-289342
	325557	6056302	Plus	50921-51050
80	325559	6249595	Minus	118590-119172
	325560	6249595	Minus	133794-133981
	325569	6249599	Plus	79927-80217
	325587	6682462	Plus	126724-126967
	325585	6682462	Plus	73476-73574
85	325597	5866992	Plus	1065020-1065089
	325639	5867002	Plus	253525-253608

WO 02/086443

	325739	5867038	Minus	205138-205269
	325740	5867038	Minus	207533-207690
	325792	6469828	Minus	1018-1176
5	325735	6552447	Minus	269122-269190
	325685	6682468	Plus	117397-117483
	325686	6682468	Plus	118337-118439
	325819	6682490	Minus	130314-130370
	329764	6048195	Minus	109733-109958
10	329703	6065793	Minus	139994-140138
	329643	6448539	Plus	53403-53537
	329816	6524888	Minus	70296-70423
	329860	6687260	Minus	163474-163605
	325883	5867087	Plus	22498-22663
15	325895	5867097	Plus	358317-358476
	325925	5867124	Plus	115749-115962
	325932	5867127	Plus	7369-7441
	325941	5867133	Minus	64228-64402
	325969	5867153	Plus	101911-102081
	325971	5867153	Plus	105841-105035
20	329993	4567166	Minus	101307-101434
	330020	6671887	Plus	172397-172491
	326163	5867168	Minus	7831-8035
	326274	5867171	Minus	410289-410404
25	326025	5867178	Plus	70854-70915
	326046	5867182	Minus	62668-62825
	326099	5867185	Minus	661381-661510
	326108	5867187	Minus	23784-23903
	326165	5867208	Minus	62787-62929
30	326189	5867212	Plus	69288-69413
	326204	5867218	Minus	148088-148200
	326230	5867230	Minus	301868-301972
	330052	4567182	Plus	352560-352963
35	330036	6042048	Plus	117120-117216
	326360	5867293	Plus	13627-13844
	326589	5867320	Plus	22760-22919
	326393	5867341	Plus	41702-41841
	326505	5867435	Minus	8818-8949
	326515	5867439	Plus	36683-36809
40	326592	6138928	Plus	23689-23828
	330107	6015249	Minus	100091-100282
	330106	6015249	Minus	99443-99778
	330100	6015253	Plus	21166-21301
	330093	6015278	Plus	1043-1199
45	330088	6015293	Plus	37517-37638
	330085	6015302	Minus	59613-59770
	330120	6671864	Minus	127553-127656
	330123	6671869	Minus	35311-35406
	326742	5867611	Minus	95187-95248
50	326605	5867637	Plus	24656-24749
	326818	6117831	Minus	15199-15309
	326720	6552456	Plus	84525-84677
	326770	6598307	Minus	513603-513668
	326692	6682502	Plus	117697-117899
55	326693	6682502	Minus	335002-335095
	326983	5867657	Minus	16023-16581
	326991	5867660	Plus	18147-18339
	326936	6004446	Minus	10217-10357
	326964	6469836	Plus	75340-75456
60	327040	6531965	Plus	783670-783817
	327053	6531965	Plus	2247267-2247437
	327075	6531965	Plus	4041318-4041431
	327085	6531965	Plus	4734947-4735069
	327036	6531965	Plus	319951-320040
65	327130	6531976	Plus	20247-22343
	327156	5866841	Minus	2462-2620
	327268	5867481	Plus	48583-48773
	327332	5867516	Minus	56361-56532
	327220	5867525	Minus	65701-65781
70	327224	5867534	Plus	188468-188544
	327321	6249562	Minus	99745-99836
	327381	6552412	Minus	61013-62130
	327396	5867743	Plus	8702-8820
	327414	5867750	Plus	102461-102586
75	327442	5867759	Plus	111483-111618
	327467	5867772	Plus	88030-88151
	327473	5867775	Plus	75101-75181
	327483	5867783	Plus	181673-181662
	327377	5867793	Minus	37610-37676
80	327562	5867804	Minus	343989-344474
	327568	5867811	Minus	46152-46287
	327606	6004463	Plus	200262-200495
	327611	5867863	Minus	175063-175392
	327642	5867891	Minus	2513-2743
85	327654	5867910	Minus	97564-97710
	327734	5867940	Minus	31003-31583

WO 02/086443

PCT/US02/12476

	327776	5867964	Minus	130791-130871
	327796	5867982	Plus	85267-85405
	327840	6249578	Minus	73065-73206
5	330208	6013599	Plus	66517-66931
	330263	6571884	Minus	101503-101634
	328004	5867993	Minus	157407-157887
	328101	5868020	Plus	289920-290014
	328100	5868020	Minus	263545-263635
10	328113	5868024	Minus	80378-80491
	328157	5868064	Plus	73326-73615
	328196	5868080	Minus	16551-16729
	328197	5868081	Minus	42133-42438
	327940	5868197	Minus	95240-95428
15	327984	5868216	Plus	66611-66677
	328021	5902482	Plus	713478-714590
	328068	6117819	Plus	253903-254022
	328264	6381912	Plus	55086-55404
	330300	2905862	Minus	3246-3302
20	328608	5868222	Minus	87770-87953
	328600	5868229	Minus	38889-40010
	328616	5868239	Plus	293920-294224
	328623	5868246	Minus	120020-120126
	328632	5868247	Plus	76734-76853
	328666	5868254	Minus	778-801
25	328698	5868264	Minus	625555-625633
	328700	5868264	Plus	764089-764203
	328708	5868271	Minus	68114-68854
	328735	5868289	Plus	89389-89455
30	328743	5868289	Plus	274638-274726
	328806	5868324	Plus	29408-29684
	328299	5868366	Minus	149708-149889
	328342	5868383	Plus	59955-60094
	328365	5868387	Minus	270724-270798
35	328369	5868388	Plus	75371-75583
	328381	5868392	Plus	662758-662848
	328451	5868425	Minus	217275-217336
	328481	5868449	Minus	8987-9180
	328500	5868464	Plus	59098-59481
40	328530	5868482	Plus	334973-335406
	328664	6004473	Plus	1193739-1193866
	328861	6381928	Minus	108317-108403
	328908	5868493	Plus	117002-117059
	328933	5868500	Plus	771755-771869
45	328934	5868500	Plus	846342-846448
	328949	6456765	Minus	43552-43619
	330313	6042030	Minus	33642-33775
	329005	5868542	Plus	85470-85673
	330366	2944106	Plus	151837-151914
50	330372	6580495	Minus	317461-317688
	329033	5868561	Minus	5390-5479
	329037	5868562	Minus	32466-32562
	329067	5868591	Minus	146417-147652
	329134	5868679	Plus	29969-30018
55	329157	5868687	Minus	145940-146155
	329178	5868704	Plus	178177-179463
	329192	5868716	Plus	166936-167020
	329194	5868716	Minus	304450-304559
	329204	5868720	Minus	3050-3190
60	329224	5868728	Plus	27422-27664
	329228	5868728	Minus	50118-50287
	329268	5868771	Plus	25554-26289
	329337	5868806	Minus	467155-467222
	329011	6682532	Plus	48658-48741

TABLE 9A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer

Table 9A shows about 1312 genes up-regulated in lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) relative to normal body tissues. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 9B show the accession numbers for those Pkey's lacking UnigenelD's for table 9A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 9C show the genomic positioning for those Pkey's lacking Unigena ID's and accession numbers in table 9A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAcct: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigena number  
 Unigena Title: Unigena gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAcct	UnigenelD	Unigena Title	R1	R2
400195			NM_007057:Homo sapiens ZW10 interactor	1.00	1.00
400205			NM_006265:Homo sapiens RAD21 (S. pombe)	15.80	396.00
400220			Eos Control	2.28	2.84
400277			Eos Control	7.68	9.72
400285			Eos Control	1.00	1.00
400288	X06256	Hs.149609	integrin, alpha 5 (fibronectin receptor,	1.04	2.24
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.85	74.00
400301	X03635	Hs.1657	estrogen receptor 1	1.00	1.00
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	1.75	1.65
400328	X87344	Hs.180062	transporter 2, ATP-binding cassette, sub	0.87	1.80
400419	AF084545		Target	156.55	253.00
400512			NM_030878:Homo sapiens cytochrome P450,	1.00	2.00
400517	AF242388		lengsin	3.67	87.00
400560			NM_030878:Homo sapiens cytochrome P450,	1.00	1.00
400664			NM_002425:Homo sapiens matrix metallopro	20.26	45.00
400665			NM_002425:Homo sapiens matrix metallopro	1.36	1.07
400666			NM_002425:Homo sapiens matrix metallopro	3.26	3.22
400749			NM_003105:Homo sapiens sortilin-related	1.00	91.00
400763			Target Exon	7.63	24.00
401027			Target Exon	1.00	1.00
401093			C12000586:gij6330167 dbj BAA85477.1  (A	1.00	155.00
401203			Target Exon	1.00	86.00
401212			C12000457:gij7512178 pir T30337 polypr	1.00	400.00
401411			ENSP00000247172:HYPOTHETICAL 126.2 kDa	1.00	72.00
401435			C14000397:gij7499898 pir T33295 hypoth	1.00	64.00
401464	AF039241		histone deacetylase 5	3.82	49.00
401714			ENSP00000241802:CDNA FLJ11007 FIS, CLON	2.02	40.00
401747			Homo sapiens keratin 17 (KRT17)	128.43	68.00
401760			Target Exon	1.74	35.00
401780			NM_005557:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275:Homo sapiens keratin 15 (KRT1	4.13	2.70
401797			Target Exon	1.44	2.10
401961			NM_021626:Homo sapiens serine carboxypep	1.41	1.86
401985	AF053004		class I cytokine receptor	1.00	177.00
401994			Target Exon	61.84	47.00
402075			ENSP00000251056:Plasma membrane calcium	1.00	1.00
402260			NM_001436:Homo sapiens fibrillarin (FBL	1.58	1.39
402265			Target Exon	2.09	35.00
402297			Target Exon	1.00	92.00
402408			NM_030920:Homo sapiens hypothetical pro	28.87	13.00
402420			C1000823:gij10432400 emb CAC10290.1  (A	1.00	1.44
402674			Target Exon	7.44	243.00
402802			NM_001397:Homo sapiens endothelin conver	1.00	70.00
402994			NM_002463:Homo sapiens myxovirus (influ	1.37	1.43
403137			NM_005391:Homo sapiens nucleolin (NCL),	1.00	19.00
403306	NM_006825		transmembrane protein (63kD), endoplasmic	1.00	43.00
403329			Target Exon	1.00	61.00
403381			ENSP00000231844:Ecotropic virus integra	1.00	119.00
403478			NM_022342:Homo sapiens kinesin protein 9	28.13	136.00
403485			C3001813:gij12737279 ref XP_012163.1  k	20.23	76.00
403627			Target Exon	6.30	29.33
403715			Target Exon	1.30	35.00
404044			ENSP00000237855:DJ398G3.2 (NOVEL PROTEI	1.00	54.00
404076			NM_016020:Homo sapiens CGI-75 protein (	14.29	91.00
404101			C8000950:gij423560 pir A47318 RNA-bind	1.00	1.00
404140			NM_006510:Homo sapiens ret finger protei	1.42	1.44
404165			ENSP00000244562:NRH dehydrogenase [quino	1.00	54.00
404185			Target Exon	1.00	117.00
404210			NM_005936:Homo sapiens myeloid/lymphoid	5.93	13.77
404253			NM_021058:Homo sapiens H2B histone fami	1.00	1.00

404287			C6001909:gi 704441 kb BAA18909.1  [D298	29.71	42.00
404298			C6001238:gi 121715 sp P26697 GTA3_CHICK	1.30	1.00
404347			Target Exon	1.00	1.00
404440			NM_021048:Homo sapiens melanoma antigen,	1.00	15.00
404721			NM_005596:Homo sapiens nuclear factor I	1.00	60.00
404794	NM_000078		cholesterol ester transfer protein, plas	1.07	1.38
404854			Target Exon	1.81	2.01
404877			NM_005365:Homo sapiens melanoma antigen,	1.00	1.00
404927			Target Exon	1.00	1.00
404996			Target Exon	1.00	1.00
405449			CY000047:gi 11427234 ref XP_009399.1  z	1.00	1.00
405568			NM_031413:Homo sapiens cat eye syndrome	1.00	78.00
405572			Target Exon	0.76	1.14
405646			C12000200:gi 4557225 ref NP_000005.1  al	1.01	1.28
405676	BE335714		cytochrome c-1	1.13	2.89
405770			NM_002362:Homo sapiens melanoma antigen,	45.52	37.00
405932			C15000305:gi 3806122 gb AAC69198.1  (AFD	1.99	1.99
406137			NM_000179:Homo sapiens mutS (E. coli) h	2.77	2.38
406360			Target Exon	1.00	35.00
406399			NM_003122:Homo sapiens serine protease	1.00	39.00
406467			Target Exon	1.00	1.00
406621	X57809	Hs.181125	immunoglobulin lambda locus	1.41	1.74
406642	AJ245210		gb:Homo sapiens mRNA for immunoglobulin	2.16	3.91
406663	U24683	Hs.293441	immunoglobulin heavy constant mu	2.07	2.93
406671	AA129547	Hs.285754	met proto-oncogene (hepatocyte growth fa	15.00	51.00
406673	M34996	Hs.198253	major histocompatibility complex, class	0.98	3.09
406676	X58399	Hs.81221	Human L2-9 transcript of unrearranged im,	1.30	1.53
406678	U77534		gb:Human clone 1A11 Immunoglobulin varia	1.33	1.45
406685	M18728		gb:Human nonspecific crossreacting anti	1.46	2.85
406687	M31126	Hs.272822	pregnancy specific beta-1-glycoprotein 9	8.61	8.50
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
406698	X03068	Hs.73931	major histocompatibility complex, class	1.01	2.52
406815	AA833930	Hs.288036	RNA isopentenylpyrophosphate transferas	20.25	32.00
406851	AA609784		major histocompatibility complex, class	0.75	1.91
406964	M21305		gb:Human alpha satellite and satellite 3	38.15	1114.00
406967	M24349		gb:Human parathyroid hormone-like protei	1.00	1.00
406974	M57293		gb:Human parathyroid hormone-related pep	1.00	1.00
407103	AA424881	Hs.256301	hypothetical protein MGC13170	1.77	1.10
407128	R83312	Hs.237260	EST	1.00	1.00
407137	T97307		gb:ye53h05.s1 Soares fetal liver spleen	142.70	135.00
407168	R45175	Hs.117183	ESTs	2.16	18.00
407239	AA076350	Hs.67846	leukocyte immunoglobulin-like receptor,	1.10	1.57
407242	M18728		gb:Human nonspecific crossreacting anti	1.12	2.85
407244	M10014	Hs.75431	fibrinogen, gamma polypeptide	3.24	15.38
407289	AA135159	Hs.203349	Homo sapiens cDNA FLJ12149 fis, clone MA	3.53	3.68
407300	AA102616	Hs.120769	gb:zn43e07.s1 Stratagene HeLa cell s3 93	19.74	73.00
407366	AF026942	Hs.271530	gb:Homo sapiens cig33 mRNA, partial sequ	0.06	8.25
407378	AA299264	Hs.57776	ESTs, Moderately similar to I38022 hypot	1.00	26.00
407430	AF169351		gb:Homo sapiens protein tyrosine phospho	1.00	25.00
407453	AJ132087		gb:Homo sapiens mRNA for axonemal dynein	1.00	75.00
407577	AW131324	Hs.246759	hypothetical protein MGC12538	1.00	1.00
407634	AW016569	Hs.136414	UDP-GlcNAc:betaGal beta-1,3-N-acetylgluc	111.20	228.00
407710	AW022727	Hs.23616	ESTs	1.00	28.00
407720	AB037776	Hs.38002	KIAA1355 protein	1.89	1.31
407746	AK001962		hypothetical protein FLJ11100	1.00	1.00
407756	AA116021	Hs.38260	ubiquitin specific protease 18	4.51	5.00
407758	D50915	Hs.38365	KIAA0125 gene product	1.00	28.00
407782	AA608956	Hs.112619	ESTs, Moderately similar to PURKINJE CEL	0.97	1.14
407788	BE514982	Hs.38991	S100 calcium-binding protein A2	7.88	3.83
407790	AJ027274	Hs.288941	Homo sapiens cDNA FLJ14866 fis, clone PL	3.63	42.00
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	89.96	109.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
407944	R34008	Hs.239727	desmocollin 2	111.30	70.00
408000	L11690	Hs.620	bulbos pemphigoid antigen 1 (230/240kD)	151.17	8.00
408031	AA081395	Hs.42173	Homo sapiens cDNA FLJ10366 fis, clone NT	9.91	93.00
408063	BE086548	Hs.42346	calcineurin-binding protein calsarcin-1	195.78	231.00
408070	AW148852		gb:xf05d05.x1 NCI_CGAP_Bm35 Homo sapien	1.00	1.00
408101	AW968504	Hs.123073	CDC2-related protein kinase 7	37.84	61.00
408122	AI432652	Hs.42824	hypothetical protein FLJ10718	0.85	1.71
408212	AA297567	Hs.43728	hypothetical protein	5.88	7.91
408243	Y00787	Hs.624	interleukin 8	4.27	9.98
408349	BE546947	Hs.44276	homeo box C10	3.79	3.46
408353	BE439838	Hs.44298	mitochondrial ribosomal protein S17	1.88	1.65
408354	AJ382803	Hs.159235	ESTs	1.00	73.00
408369	R38438	Hs.182575	solute carrier family 15 (H777) transport	1.41	16.50
408380	AF123050	Hs.44532	diubiquitin	15.19	37.22
408482	NM_000676	Hs.45743	adenosine A2b receptor	1.65	1.19
408522	AJ541214	Hs.46320	Small proline-rich protein SPRK (human,	1.98	1.24
408536	AW381532	Hs.135188	ESTs	1.55	1.50
408545	AW235405	Hs.253690	ESTs	1.00	1.00
408572	AA055611	Hs.226568	ESTs, Moderately similar to ALU4_HUMAN A	1.00	44.00
408633	AW963372	Hs.46677	PRO2000 protein	107.16	56.00
408660	AA525775		ESTs, Moderately similar to PC4259 ferri	1.00	1.00
408761	AA057264	Hs.238936	ESTs, Weakly similar to (define not ava	52.24	141.00
408771	AW732573	Hs.47584	potassium voltage-gated channel, delayed	3.05	109.00



	408783	AF192522	Hs.47701	NPC1 (Niemann-Pick disease, type C1, gen	1.02	1.07
	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	41.19	61.00
	408805	H69912	Hs.48269	vaccinia related kinase 1	24.67	45.00
5	408841	AW438865	Hs.256862	ESTs	1.00	58.00
	408873	AL046017	Hs.182278	calmodulin 2 (phosphorylase kinase, delt	1.00	89.00
	408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
	408992	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.00	1.00
	408996	AI79168	Hs.344096	glycoprotein (transmembrane) nmb	3.71	5.50
10	409015	BE389387	Hs.49767	NM_004553:Homo sapiens NADH dehydrogenas	1.44	1.24
	409038	T97490	Hs.50002	small inducible cytokine subfamily A (Cy	4.28	5.32
	409041	AB033025	Hs.50081	Hypothetical protein, XP_051850 (KIAA119	112.42	195.00
	409077	AA401369	Hs.190721	ESTs	1.00	17.00
	409093	BE243834	Hs.50441	CGI-04 protein	2.02	1.93
15	409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
	409142	AL136877	Hs.50758	SMC4 (structural maintenance of chromoso	14.87	6.00
	409187	AF154530	Hs.50966	carbamoyl-phosphate synthetase 1, mitocho	1.00	1.00
	409228	AI654298	Hs.271695	ESTs, Weakly similar to 2109260A B cell	1.22	1.00
	409234	AI879419	Hs.27206	ESTs	1.00	1.00
20	409268	AA625304	Hs.187579	ESTs	11.90	23.00
	409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
	409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	168.91	35.00
	409404	BE220053	Hs.129066	ESTs	1.00	1.00
	409420	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kallini	79.74	96.00
	409430	R21945	Hs.346735	splicing factor, arginine/serine-rich 5	1.45	2.10
25	409446	AI561173	Hs.67688	ESTs	1.00	4.00
	409506	NM_006153	Hs.54589	NCK adaptor protein 1	3.97	28.00
	409522	AA075382		gb:zm87b03.s1 Stratagene ovarian cancer	15.98	141.00
	409582	AA401369	Hs.190721	ESTs	1.00	17.00
30	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibitor	292.12	79.00
	409705	M37762	Hs.56023	brain-derived neurotrophic factor	1.00	82.00
	409719	AI769160	Hs.108681	Homo sapiens brain tumor associated prot	1.00	1.00
	409731	AA126985	Hs.56145	thymosin, beta, identified in neuroblast	0.12	18.12
	409744	AW676268	Hs.56265	Homo sapiens mRNA: cDNA DKFZp586P2321 (f	20.75	51.00
	409757	NM_001898	Hs.123114	cystatin SN	22.46	15.80
35	409866	AW502152		gb:U1-HF-BR0p-ajr-f-11-Q.U1.r1 NIH_MGC_5	1.00	1.00
	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	1.50	1.09
	409902	AI337658	Hs.156351	ESTs	25.92	50.00
	409935	AW511413	Hs.278025	ESTs	2.63	2.11
40	409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	2.17	4.01
	409958	NM_001523	Hs.57697	hyaluronan synthase 1	0.91	2.07
	410001	AB041036	Hs.57771	kallikrein 11	1.04	2.28
	410032	BE065985		gb:RC3-BT0319-120200-014-a09 BT0319 Homo	1.00	58.00
	410037	AB020725	Hs.58009	KIAA0918 protein	1.00	34.00
45	410044	BE566742	Hs.58169	highly expressed in cancer, rich in leuc	1.00	1.00
	410048	W76467	Hs.58218	proline oxidase homolog	1.03	1.44
	410076	T05387	Hs.7991	ESTs	1.12	1.50
	410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
	410153	BE311926	Hs.15830	hypothetical protein FLJ12691	1.00	1.00
50	410166	AK001376	Hs.59346	hypothetical protein FLJ10514	1.00	1.00
	410193	AJ132592	Hs.59757	zinc finger protein 281	42.01	51.00
	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	1.72	1.32
	410309	BE043077	Hs.278153	ESTs	1.00	2.00
	410340	AW182833	Hs.112188	hypothetical protein FLJ13149	32.08	75.00
55	410348	AW182663	Hs.95469	ESTs	1.00	1.00
	410407	X66839	Hs.63287	carbonic anhydrase IX	1.40	1.11
	410418	D31382	Hs.63325	transmembrane protease, serine 4	4.30	2.03
	410438	AB037756	Hs.45207	hypothetical protein KIAA1335	1.00	18.00
	410553	AW016824	Hs.255527	hypothetical protein MGC14128	1.34	1.04
60	410555	W27235	Hs.64311	a disintegrin and metalloproteinase doma	23.99	1.41
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
	410681	AW246890	Hs.65425	calbindin 1, (28kD)	10.88	18.92
	410781	AI375672	Hs.165028	ESTs	1.00	57.00
	411027	AF072099	Hs.67846	leukocyte immunoglobulin-like receptor,	1.62	3.78
65	411074	X60435	Hs.68137	adenylate cyclase activating polypeptide	1.00	1.15
	411089	AA456454		cell division cycle 2-like 1 (PITSLRE pr	1.56	1.58
	411152	BE069199		gb:QV3-BT0379-010300-105-g03 BT0379 Homo	1.00	84.00
	411248	AA551538	Hs.334605	Homo sapiens cDNA FLJ14408 fis, clone HE	1.82	1.45
	411252	AB018549	Hs.69328	MD-2 protein	7.32	12.74
70	411263	BE297802	Hs.69360	kinasin-like 6 (mitotic centromere-assoc	3.44	2.55
	411365	M76477	Hs.289082	GM2 ganglioside activator protein	1.35	2.02
	411402	BE297855	Hs.69855	NRAS-related gene	1.00	46.00
	411573	AB029000	Hs.70823	KIAA1077 protein	11.40	11.35
	411579	AC005258	Hs.70830	U6 snRNA-associated Sm-like protein LSM7	1.08	1.90
75	411617	AA247994	Hs.90063	neurocalcin delta	1.74	2.57
	411732	AA059325	Hs.71642	guanine nucleotide binding protein (G pr	1.02	1.00
	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	1.34	2.19
	411789	AF245505	Hs.72157	Adlcan	2.19	2.79
	411800	N39342	Hs.103042	microtubule-associated protein 1B	23.34	34.00
80	411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	1.00	8.00
	412115	AK001763	Hs.73239	hypothetical protein FLJ10901	2.07	1.64
	412140	AA219691	Hs.73825	RAB6 interacting, kinesin-like (rakines	118.48	92.00
	412276	BE262621	Hs.73798	macrophage migration inhibitory factor (	1.98	1.49
	412464	T78141	Hs.22826	ESTs, Weakly similar to I55214 salivary	1.16	1.34
85	412530	AA766268	Hs.266273	hypothetical protein FLJ13346	41.52	84.00
	412537	AL031778		nuclear transcription factor Y, alpha	17.90	55.00

	412659	AW753865	Hs.74376	olfactomedin related ER localized protel	14.65	47.00
	412719	AW016610	Hs.816	ESTs	382.46	128.00
	412723	AA648459	Hs.335951	hypothetical protein AF301222	54.90	1.00
	412811	H06382		ESTs	1.00	11.00
5	412817	AL037159	Hs.74619	proteasome (prosome, macropain) 26S subu	1.63	1.42
	412863	AA121673	Hs.59757	zinc finger protein 281	17.63	56.00
	412924	BE018422	Hs.75258	H2A histone family, member Y	1.00	22.00
	413004	T35501	Hs.75117	interleukin enhancer binding factor 2, 4	2.19	2.05
10	413011	AW068115	Hs.821	biglycan	1.22	1.88
	413048	M93221	Hs.75182	mannose receptor, C type 1	0.30	6.23
	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote	3.43	8.71
	413129	AF292100	Hs.104613	RP42 homolog	4.67	4.77
	413142	M81740	Hs.75212	ornithine decarboxylase 1	1.92	2.59
15	413223	AI732182	Hs.191866	ESTs	5.73	27.00
	413248	T64858	Hs.21433	hypothetical protein DKFZp547J036	0.99	1.06
	413273	U75679	Hs.75257	stem-loop (histone) binding protein	1.00	18.00
	413278	BE563085	Hs.833	interferon-stimulated protein, 15 kDa	1.10	1.09
	413281	AA861271	Hs.222024	transcription factor BMAL2	95.94	69.00
20	413364	BE536218	Hs.137516	fidgulin-like 1	1.00	1.00
	413385	M34455	Hs.840	indoleamine-pyrole 2,3 dioxygenase	0.95	2.09
	413409	AI638418	Hs.1440	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	1.00	1.00
	413453	AA129640	Hs.128065	ESTs	1.00	31.00
	413527	BE250788	Hs.179882	hypothetical protein FLJ12443	1.08	1.46
	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	79.15	114.00
25	413573	AI733859	Hs.149089	ESTs	1.00	1.00
	413582	AW295647	Hs.71331	hypothetical protein MGC5350	8.80	10.00
	413597	AW302885	Hs.117183	ESTs	1.00	1.00
	413690	BE157489		gb:RC1-HT0375-120200-011-e06 HT0375 Homo	1.00	1.00
30	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B	3.16	2.32
	413719	BE439580	Hs.75498	small inducible cytokine subfamily A (Cy	2.88	9.52
	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kalinin	144.10	108.00
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	1.00	17.00
	413833	Z15005	Hs.75573	centromere protein E (312kD)	1.00	1.00
35	413882	AA132973	Hs.184492	ESTs	64.24	148.00
	413926	AA133338	Hs.54310	ESTs	1.00	67.00
	413943	AW294416	Hs.144687	Homo sapiens cDNA FLJ12981 fis, clone NT	43.42	42.00
	413995	BE048146	Hs.75571	syntaxin 1A (brain)	1.23	1.11
	414035	Y00630	Hs.75716	serine (or cysteine) proteinase inhibito	2.02	2.51
40	414142	AW368397	Hs.334485	Homo sapiens cDNA FLJ14438 fis, clone HE	1.00	102.00
	414180	AI863304	Hs.120905	Homo sapiens cDNA FLJ11448 fis, clone HE	6.92	77.00
	414245	BE148072	Hs.75850	WAS protein family, member 1	1.00	1.00
	414275	AW970254	Hs.889	Charot-Leyden crystal protein	1.00	59.00
	414317	BE263280	Hs.75888	phosphogluconate dehydrogenase	1.52	1.73
45	414334	AA824298	Hs.21331	hypothetical protein FLJ10036	1.78	1.72
	414341	D80004	Hs.75909	KIAA0182 protein	33.90	151.00
	414368	W70171	Hs.75939	uridine monophosphate kinase	171.60	97.00
	414416	AW409885	Hs.76084	hypothetical protein MGC2721	2.32	1.85
	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1	226.15	66.00
50	414570	Y00285	Hs.76473	insulin-like growth factor 2 receptor	1.64	1.98
	414618	AI204600	Hs.66978	hypothetical protein MGC10764	1.87	72.00
	414675	R79015	Hs.296281	interleukin enhancer binding factor 1	1.51	1.39
	414683	S78296	Hs.76688	hypothetical protein MGC12702	43.61	64.00
	414696	AF002020	Hs.76918	Niemann-Pick disease, type C1	28.63	71.00
55	414711	AI310440	Hs.288735	Homo sapiens cDNA FLJ13522 fis, clone PL	14.86	42.00
	414718	H95348	Hs.107987	ESTs	1.00	5.00
	414732	AW410976	Hs.77152	minichromosome maintenance deficient (S.	1.64	1.44
	414747	U30872	Hs.77204	centromere protein F (350/400kD, mitosi	65.01	74.00
	414761	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2	130.35	121.00
60	414774	X02419	Hs.77274	plasminogen activator, urokinase	2.24	2.19
	414806	D14694	Hs.77329	phosphatidylserine synthase 1	1.63	1.53
	414809	AI434699	Hs.77356	transferrin receptor (p90, CD71)	1.97	2.60
	414812	X72755	Hs.77367	monokine induced by gamma Interferon	3.48	10.60
	414825	X06370	Hs.77432	epidermal growth factor receptor (avian	103.22	143.00
65	414839	X63692	Hs.77462	DNA (cytosine-5)-methyltransferase 1	1.80	1.69
	414883	AA926960		CDC28 protein kinase 1	14.29	10.06
	414907	X90725	Hs.77597	polo (Drosophila)-like kinase	1.95	2.20
	414914	U49844	Hs.77613	ataxia telangiectasia and Rad3 related	3.00	2.90
	414945	BE076358	Hs.77667	lymphocyte antigen 6 complex, locus E	1.02	1.21
70	414972	BE263782	Hs.77695	KIAA0008 gene product	1.00	1.00
	415014	AW954064	Hs.24951	ESTs	1.42	2.84
	415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	34.72	107.00
	415227	AW821113	Hs.72402	ESTs	1.87	49.00
75	415238	R37780	Hs.21422	ESTs	1.00	1.00
	415263	AA948033	Hs.130853	ESTs	1.00	1.00
	415295	R41450	Hs.6546	ESTs	1.00	1.00
	415339	NM_015156	Hs.78398	KIAA0071 protein	51.18	166.00
	415669	NM_006025	Hs.78589	serine (or cysteine) proteinase inhibito	30.84	63.00
80	415674	BE394784	Hs.78596	proteasome (prosome, macropain) subunit,	1.48	1.39
	415709	AA649850	Hs.278558	ESTs	1.00	1.00
	415735	AA704162	Hs.120811	ESTs, Weakly similar to I38022 hypotheti	1.00	72.00
	415799	AA653718	Hs.225841	DKFZP434D193 protein	6.23	31.00
	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
85	415857	AA866115	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	32.51	35.00
	415989	AI267700		ESTs	78.89	1.00

	416018	AW138239	Hs.78977	proprotein convertase subtilisin/kexin 1	1.00	1.00
	416065	BE267931	Hs.78996	proliferating cell nuclear antigen	3.35	2.32
	416111	AA033813	Hs.78018	chromatin assembly factor 1, subunit A (	39.03	3.00
5	416177	AA174069	Hs.187607	ESTs	1.00	9.00
	416178	AI808527	Hs.192822	serologically defined breast cancer anti	3.83	3.76
	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h	9.70	1.00
	416239	AL038450	Hs.48948	ESTs	83.87	129.00
10	416250	AA581386	Hs.73452	hypothetical protein MGC10791	1.96	2.12
	416322	BE019494	Hs.79217	pyrroline-5-carboxylate reductase 1	2.08	1.73
	416423	H54375	Hs.268921	ESTs	1.00	89.00
	416448	L13210	Hs.79339	lectin, galactoside-binding, soluble, 3	1.28	1.54
	416498	U33632	Hs.79351	potassium channel, subfamily K, member 1	27.29	67.00
	416658	U03272	Hs.79432	fibrillin 2 (congenital contractual era	53.29	51.00
15	416661	AA634543	Hs.79440	IGF-II mRNA-binding protein 3	9.96	5.00
	416722	AA354604	Hs.122546	hypothetical protein FLJ23017	3.68	33.00
	416819	U77735	Hs.80205	pim-2 oncogene	1.59	1.84
	416936	N21352	Hs.42987	ESTs, Weakly similar to S21348 probable	1.00	1.00
20	417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	32.95	156.00
	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	3.91	4.93
	417218	AA129547	Hs.285754	met proto-oncogene (hepatocyte growth fa	1.00	51.00
	417233	W25005	Hs.24395	small inducible cytokine subfamily B (Cy	3.38	2.05
	417308	H60720	Hs.81892	KIAA0101 gene product	82.94	25.35
25	417315	AI080042	Hs.180450	ribosomal protein S24	106.61	121.00
	417324	AW265494		ESTs	1.20	1.28
	417366	BE185289	Hs.1076	small proline-rich protein 1B (cornifin)	8.97	3.27
	417389	BE260954	Hs.82045	midkine (neurite growth-promoting factor	2.69	1.82
30	417428	N87578	Hs.278871	gb:LL2030F Human fetal heart, Lambda ZAP	1.00	52.00
	417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein	304.75	173.00
	417466	AI681547	Hs.59457	hypothetical protein FLJ22127	1.24	1.34
	417512	AI979168	Hs.344096	glycoprotein (transmembrane) nmb	2.14	5.50
	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated	2.66	1.68
35	417542	JO4129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
	417576	AA339449	Hs.82285	phosphoribosylglycinamide formyltransfer	42.76	51.00
	417715	AW969587	Hs.86366	ESTs	6.35	2.75
	417720	AA205625	Hs.208067	ESTs	113.31	58.00
	417791	AW865339	Hs.111471	ESTs	39.98	16.00
40	417830	AW504786	Hs.122579	hypothetical protein FLJ10461	2.61	31.00
	417866	AW067503	Hs.82772	collagen, type XI, alpha 1	2.35	2.44
	417900	BE250127	Hs.82906	CDC20 (cell division cycle 20, S. cerevi	1.52	1.11
	417933	X02308	Hs.82962	thymidylate synthetase	4.74	2.55
	417944	AU077196	Hs.82985	collagen, type V, alpha 2	3.61	5.21
45	417975	AA641836	Hs.30085	hypothetical protein FLJ23186	12.49	38.00
	417991	AA731452	Hs.190008	ESTs	1.00	26.00
	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member	3.02	2.12
	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	187.59	1.00
	418054	NM_002318	Hs.83354	lysyl oxidase-like 2	2.85	2.63
50	418057	NM_012151	Hs.83363	coagulation factor VIII-associated (inir	1.54	1.69
	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4	6.82	5.22
	418140	BE613636	Hs.83551	microfibrillar-associated protein 2	1.26	1.46
	418203	X54942	Hs.83758	CDC28 protein kinase 2	134.19	144.00
	418207	C14685	Hs.34772	ESTs	1.00	1.00
55	418216	AA662240	Hs.283099	AF15q14 protein	64.66	61.00
	418236	AW994005	Hs.337534	ESTs	18.53	147.00
	418249	H89226	Hs.34892	KQA1323 protein	30.53	106.00
	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	1.00	3.00
	418283	S79895	Hs.83942	calhepsin K (pseudosynostosis)	3.96	5.16
60	418300	AI433074	Hs.86682	Homo sapiens cDNA: FLJ21578 fis, clone C	3.18	2.91
	418322	AA284166	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	11.96	6.68
	418327	U70370	Hs.84136	paired-like homeodomain transcription fa	9.23	2.22
	418345	AJ001696	Hs.241407	serine (or cysteine) proteinase inhibito	1.00	1.00
	418379	AA218940	Hs.137516	fidgetin-like 1	21.68	44.00
65	418397	NM_001269	Hs.84746	chromosome condensation 1	1.00	8.00
	418403	D86978	Hs.84790	KQA0225 protein	16.91	18.98
	418462	BE001596	Hs.85266	Integrin, beta 4	1.56	1.16
	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	3.22	2.38
	418506	AA084248	Hs.85339	G protein-coupled receptor 39	2.66	2.22
70	418526	BE019020	Hs.85838	solute carrier family 16 (monocarboxylic	2.04	2.21
	418538	BE244323	Hs.85951	exportin, tRNA (nuclear export receptor	1.33	37.00
	418543	NM_005329	Hs.85962	hyaluronan synthase 3	1.04	1.23
	418574	N28754		M-phase phosphoprotein 9	48.50	85.00
	418592	X99226	Hs.284153	Fanconi anemia, complementation group A	18.24	26.00
75	418641	BE243136	Hs.86947	a disintegrin and metalloproteinase doma	1.19	1.41
	418661	NM_001949	Hs.1189	E2F transcription factor 3	29.05	43.00
	418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
	418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
	418686	Z36830	Hs.87268	annexin A8	1.54	1.98
80	418689	AI360883	Hs.274448	hypothetical protein FLJ11029	1.19	1.04
	418712	Z42183		gb:HSC0BF041 normalized infant brain cDN	1.00	12.00
	418727	AA227609	Hs.94834	ESTs	1.00	49.00
	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino	49.85	1.00
	418819	AA228776	Hs.191721	ESTs	1.00	140.00
85	418830	BE513731	Hs.88959	hypothetical protein MGC4816	20.97	23.00
	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	57.09	35.00

	418971	AA360392	Hs.87113	ESTs	1.00	12.00
	418973	AA233056	Hs.191518	ESTs	4.89	28.00
	419078	M93119	Hs.89584	insulinoma-associated 1	1.00	10.00
5	419079	AW014836	Hs.18844	ESTs	1.09	1.98
	419080	AW150835	Hs.18878	hypothetical protein FLJ21620	2.06	1.68
	419088	AI538323	Hs.52620	integrin, beta 8	15.60	51.00
	419092	J05581	Hs.89603	mucin 1, transmembrane	1.11	1.83
	419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
10	419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor I	1.10	1.14
	419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	3.18	2.43
	419288	AA256106	Hs.87507	ESTs	1.00	34.00
	419335	AW960146	Hs.284137	hypothetical protein FLJ12888	1.00	8.00
15	419354	M62839	Hs.1252	apolipoprotein H (beta-2-glycoprotein I)	22.63	54.00
	419359	AL043202	Hs.90073	chromosome segregation 1 (yeast homolog)	2.50	1.98
	419423	D26488	Hs.90315	KIAA0007 protein	1.00	7.00
	419443	D62703		gb:HUM316G10B Clontech human aorta polyA	1.00	12.00
	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	1.64	1.84
20	419474	AW968619	Hs.155849	ESTs	13.63	62.00
	419485	AA489023	Hs.99807	ESTs, Weakly similar to unnamed protein	4.27	2.26
	419488	AA316241	Hs.90691	nucleophosmin/nucleoplasmin 3	3.66	3.63
	419502	AU076704		fibrinogen, A alpha polypeptide	13.05	115.00
	419539	AF070590	Hs.90869	Homo sapiens clones 24622 and 24623 mRNA	74.60	117.00
25	419556	U29615	Hs.91093	chitinase 1 (chitotriosidase)	1.47	4.98
	419569	AI971651	Hs.91143	jagged 1 (Alagille syndrome)	1.00	4.00
	419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	94.30	94.00
	419703	AI793257	Hs.128151	ESTs	15.26	50.00
	419721	NM_001650	Hs.288650	aquaporin 4	1.00	191.00
30	419729	AA586442	Hs.21411	gb:nc53a03.s1 NCL CGAP SS1 Homo sapiens	59.00	59.00
	419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	2.02	1.08
	419745	AF042001	Hs.93005	slug (chicken homolog), zinc finger prot	1.00	1.00
	419752	AA249573	Hs.152618	ESTs, Moderately similar to ZN91_HUMAN Z	29.87	77.00
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	50.99	214.00
35	419936	AI792788		gb:ol91d05.y5 NCL CGAP_Kid5 Homo sapiens	1.00	1.00
	419937	AB040959	Hs.93836	DKFZP434N014 protein	1.64	2.47
	419983	W55956	Hs.94030	Homo sapiens mRNA; cDNA DKFZp586E1624 (f	15.72	94.00
	420005	AW271106	Hs.133294	ESTs	3.15	1.43
	420047	AI478658	Hs.94631	brafeldin A-inhibited guanine nucleotide	12.45	39.00
40	420058	AK001423	Hs.94694	Homo sapiens cDNA FLJ10561 fis, clone NT	1.00	117.00
	420162	BE378432	Hs.95577	cyclin-dependent kinase 4	1.43	1.21
	420251	AW374968	Hs.348112	Human DNA sequence from clone RP5-1103G7	2.35	3.23
	420259	AF004884	Hs.96253	calcium channel, voltage-dependent, P/Q	0.77	1.15
	420281	AI623693	Hs.323494	ESTs	45.04	54.00
45	420309	AW043637	Hs.21766	ESTs, Weakly similar to ALU5_HUMAN ALU S	49.22	31.00
	420332	NM_001756	Hs.1305	serine (or cysteine) proteinase inhibitor	0.05	2.82
	420380	AA640891	Hs.102406	ESTs	0.99	2.74
	420462	AF060147	Hs.97932	chondromodulin I precursor	1.00	1.00
	420520	AK001978	Hs.98510	similar to rab11-binding protein	49.74	133.00
50	420552	AK000492	Hs.98806	hypothetical protein	94.65	88.00
	420560	AW207748	Hs.59115	ESTs	1.00	17.00
	420610	AI683183	Hs.99348	distal-less homeo box 5	1.00	13.00
	420689	H79979	Hs.88678	ESTs	50.09	95.00
	420721	AA927802	Hs.159471	ZAP3 protein	1.00	31.00
55	420759	T11832	Hs.127797	Homo sapiens cDNA FLJ11381 fis, clone HE	1.00	48.00
	420783	AI659838	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
	420900	AL045633	Hs.44269	ESTs	2.24	7.00
	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
	421002	AF116030	Hs.100932	transcription factor 17	1.00	27.00
60	421027	AA761198	Hs.55254	ESTs	2.87	38.00
	421037	AI684808	Hs.197653	ESTs	1.00	46.00
	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	1.00	98.00
	421073	NM_004689	Hs.101448	metastasis associated 1	1.34	1.46
	421110	AJ250717	Hs.1355	cathepsin E	119.47	427.00
65	421133	AA401369	Hs.190721	ESTs	1.10	17.00
	421150	AI913662	Hs.189902	ESTs	1.45	1.63
	421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
	421307	BE539976	Hs.103305	Homo sapiens mRNA; cDNA DKFZp434B0425 (f	1.37	1.10
	421316	AA287203	Hs.324728	SMA5	1.00	21.00
70	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	1.92	3.94
	421451	AA291377	Hs.50831	ESTs	5.89	14.00
	421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
	421506	BE302796	Hs.105097	thymidine kinase 1, soluble	1.56	1.08
	421508	NM_004833	Hs.105115	absent in melanoma 2	5.11	5.23
75	421516	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase I, I	1.00	3.00
	421524	AA312082	Hs.105445	GDNF family receptor alpha 1	2.63	10.58
	421526	AL080121	Hs.105460	DKFZP584C0823 protein	1.46	1.88
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	30.21	50.32
	421574	AJ000152	Hs.105924	defensin, beta 2	1.67	1.74
80	421582	AI910275		trafoll factor 1 (breast cancer, estroge	1.23	1.00
	421633	AF121860	Hs.106260	sorting nexin 10	1.00	116.00
	421659	NM_014459	Hs.106511	protocadherin 17	0.05	6.33
	421677	H64092	Hs.38282	ESTs	1.31	1.42
	421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	1.41	1.20
85	421773	W69233	Hs.112457	ESTs	1.12	1.14
	421777	BE562088	Hs.108196	HSPC037 protein	1.97	1.29

	421800	AA298151	Hs.222969	ESTs	1.03	1.30
	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.88	1.59
	421896	N62293	Hs.45107	ESTs	11.84	22.80
5	421928	AF013758	Hs.109643	polyadenylate binding protein-interactin	45.89	90.00
	421931	NM_000814	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	1.13	1.49
	421948	L42583	Hs.334309	keratin 6A	51.83	20.25
	421975	AW961017	Hs.6459	hypothetical protein FLJ11856	1.17	1.15
	422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
10	422094	AF129535	Hs.272027	F-box only protein 5	67.61	62.00
	422095	AI868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
	422109	ST3265	Hs.1473	gastrin-releasing peptide	4.18	95.50
	422128	AW881145		gb:QV0-OT0033-010400-182-a07 OT0033 Homo	40.89	71.00
	422129	AU076635	Hs.1478	serine (or cysteine) proteinase inhibito	1.13	1.38
	422134	AW179019	Hs.112110	mitochondrial ribosomal protein L42	41.59	96.00
15	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias	3.29	1.68
	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	4.93	5.73
	422282	AF019225	Hs.114309	apolipoprotein L	1.49	1.71
	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.carevis	25.99	10.91
20	422310	AA316622	Hs.98370	cytochrome P450, subfamily IIS, polypept	1.54	1.41
	422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
	422330	D30783	Hs.115263	epiregulin	1.00	112.00
	422364	AF067800	Hs.115515	C-type (calcium dependent, carbohydrate-	9.39	60.00
	422406	AF025441	Hs.116206	Opa-interacting protein 5	18.33	53.00
25	422424	AI186431	Hs.296638	prostate differentiation factor	1.71	3.21
	422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
	422487	AJ010901	Hs.198267	mucin 4, tracheobronchial	73.68	35.54
	422511	AU076442	Hs.117938	collagen, type XVI, alpha 1	173.97	26.00
	422515	AW500470	Hs.117950	multifunctional polypeptide similar to S	4.68	2.92
30	422656	AI870435	Hs.1569	LIM homeobox protein 2	1.00	1.00
	422737	M26939	Hs.119571	collagen, type III, alpha 1 (Ehlers-Danl	3.89	4.55
	422756	AA441787	Hs.119689	glycoprotein hormones, alpha polypeptide	1.05	1.46
	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	3.88	1.53
	422809	AK001379	Hs.121028	hypothetical protein FLJ10549	99.56	53.00
35	422887	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
	422938	NM_001809	Hs.1594	centromere protein A (17kD)	70.46	61.00
	422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	77.74	3.00
	422960	AW890487	Hs.63984	cadherin 13, H-cadherin (heart)	5.88	8.55
	422963	AA401369	Hs.190721	ESTs	171.41	17.00
40	422976	AU076657	Hs.1600	chaperonin containing TCP1, subunit 5 (e	2.12	1.62
	422981	AF026445	Hs.122752	TATA box binding protein (TBP)-associate	10.49	35.00
	422986	AA319777	Hs.221974	ESTs	12.40	32.47
	423034	AL119930		gb:DKFZp761A092_r1 761 (synonym: hamy2)	16.41	60.00
45	423049	X59373	Hs.188023	ESTs, Moderately similar to HXDA_HUMAN H	1.00	1.00
	423081	AF262992	Hs.123159	sperm associated antigen 4	1.82	2.96
	423184	NM_004428	Hs.1624	ephrin-A1	1.14	1.53
	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys	2.14	1.69
	423248	AA380177	Hs.125845	ribulose-5-phosphate-3-epimerase	7.18	14.00
	423309	BE006775	Hs.126782	sushi-repeat protein	21.80	64.00
50	423361	AW170055	Hs.47628	ESTs	1.00	1.00
	423453	AW450737	Hs.128791	CGI-09 protein	55.52	66.00
	423511	AF036329	Hs.129715	gonadotropin-releasing hormone 2	0.88	1.17
	423516	AB007933	Hs.129729	ligand of neuronal nitric oxide synthase	1.76	5.40
	423551	AA327598	Hs.233785	ESTs	3.54	4.33
55	423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
	423575	C18863	Hs.163443	Homo sapiens cDNA FLJ11576 fs, clone HE	38.88	70.00
	423624	AI807408	Hs.166368	ESTs	1.00	67.00
	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	76.02	1.00
	423642	AW452650	Hs.157148	hypothetical protein MGC13204	19.14	58.00
60	423662	AA642452	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro	3.61	13.57
	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	240.73	40.00
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	1.00	59.00
	423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
	423761	NM_006194	Hs.132576	paired box gene 9	1.00	1.00
65	423787	AJ295745	Hs.236204	nuclear pore complex protein	7.18	6.64
	423816	AF151064		hypothetical protein	1.00	44.00
	423826	U20325	Hs.1707	cocaine- and amphetamine-regulated trans	1.00	1.00
	423849	AL157425	Hs.133315	Homo sapiens mRNA; cDNA DKFZp761J1324 (f	1.00	1.00
	423887	AL080207	Hs.134585	DKFZP434G232 protein	1.00	1.00
70	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f	31.33	31.00
	423954	AW753164	Hs.288604	KIAA1632 protein	5.81	10.87
	423961	D13666	Hs.136348	osteoblast specific factor 2 (fasciclin	3.55	3.30
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	233.42	68.00
	424016	AW163729	Hs.6140	hypothetical protein MGC15730	0.93	1.01
75	424028	AF055084	Hs.153692	Homo sapiens cDNA FLJ14354 fs, clone Y7	21.30	52.00
	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
	424086	AI351010	Hs.102267	lysyl oxidase	21.91	70.00
	424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
	424120	T80579	Hs.290279	ESTs	1.00	1.00
80	424165	AW582904	Hs.142255	islet amyloid polypeptide	1.00	34.00
	424200	AA337221		gb:EST41944 Endometrial tumor Homo sapie	13.06	48.00
	424279	L29306	Hs.171814	tryptophan hydroxylase (tryptophan 5-mon	1.00	1.00
	424308	AW975531	Hs.154443	minichromosome maintenance deficient (S	164.58	87.00
	424326	NM_014479	Hs.145296	disintegrin protease	53.72	302.00
85	424340	AA339036	Hs.7033	ESTs	0.88	1.15

	424351	BE622117	Hs.145567	hypothetical protein	0.93	1.03
	424364	AW383226	Hs.201189	ESTs, Weakly similar to G01763 atrophin-	7.02	3.24
	424381	AA285249	Hs.146329	protein kinase Chk2	95.55	92.00
5	424411	NM_005209	Hs.146549	crystallin, beta A2	1.63	3.25
	424420	BE614743	Hs.146688	prostaglandin E synthase	1.63	1.33
	424441	X14850	Hs.147097	H2A histone family, member X	1.82	1.29
	424502	AF242388	Hs.149585	lengsin	1.00	1.00
	424503	X06256	Hs.149609	Integrin, alpha 5 (fibronectin receptor,	1.02	2.24
10	424513	BE385864	Hs.149894	mitochondrial translational initiation f	1.00	17.00
	424539	L02911	Hs.150402	Activin A receptor, type I (ACVR1) (ALK	32.46	108.00
	424568	AF005418	Hs.150595	cytochrome P450, subfamily XXVIA, polype	3.40	2.58
	424602	AK002055	Hs.151046	hypothetical protein FLJ11193	31.87	25.00
	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub	3.58	2.37
15	424645	NM_014682	Hs.151449	KIAA0535 gene product	1.00	1.00
	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	2.12	2.23
	424717	AW992282	Hs.152213	wingless-type MMTV integration site fami	1.00	1.00
	424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
	424840	D79987	Hs.153479	extra spindle poles, S. cerevisiae, homo	2.65	1.30
20	424867	A1024860	Hs.153591	Not56 (D. melanogaster)-like protein	1.23	1.05
	424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
	424979	D87989	Hs.154073	UDP-galactose transporter related	1.36	1.35
	424999	AW953120		gb:EST365190 MAGB resequences, MAGB Homo	1.24	1.41
	425048	H05468	Hs.164502	ESTs	1.00	11.00
25	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol	7.46	87.00
	425081	X74794	Hs.154443	minichromosome maintenance deficient (S.	2.52	3.82
	425118	AU076611	Hs.154672	methylene tetrahydrofolate dehydrogenase	4.84	4.03
	425159	NM_004341	Hs.154868	carbamoyl-phosphate synthetase 2, aspart	3.62	2.73
	425202	AW962282	Hs.152049	ESTs, Weakly similar to I38022 hypotheti	1.00	53.00
30	425234	AW152225	Hs.165909	ESTs, Weakly similar to I38022 hypotheti	100.77	44.00
	425236	AW067800	Hs.155223	stanniocalcin 2	3.30	2.90
	425245	A1751768	Hs.155314	KIAA0095 gene product	1.91	2.32
	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	1.41	1.49
	425266	J00077	Hs.155421	alpha-fetoprotein	1.00	68.00
35	425274	BE281191	Hs.155462	minichromosome maintenance deficient (mi	1.97	1.63
	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic	141.49	123.00
	425349	AA425234	Hs.159886	ribose 5-phosphate isomerase A (ribose 5	1.00	84.00
	425371	D49441	Hs.155981	mesothelin	0.87	1.59
	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	14.90	5.76
40	425420	BE536911	Hs.234545	hypothetical protein NUF2R	1.00	1.00
	425424	NM_004954	Hs.157199	ELKL motif kinase	10.58	9.74
	425483	AF231022	Hs.158159	FAT tumor suppressor (Drosophila) homolo	1.74	1.40
	425566	AW162943	Hs.250618	UL16 binding protein 2	1.49	1.14
	425580	L11144	Hs.1907	galanin	53.29	233.00
45	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
	425692	D90041	Hs.155956	N-acetyltransferase 1 (arylamine N-acety	1.00	55.00
	425695	NM_005401	Hs.159238	protein tyrosine phosphatase, non-recept	1.00	10.00
	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg	1.00	41.00
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	1.00	48.00
50	425810	A1923627	Hs.31903	ESTs	27.39	98.00
	425811	AL039104	Hs.159557	karyopherin alpha 2 (RAG cohort 1, Impor	1.99	1.58
	425849	A1077288	Hs.296323	serum/glucocorticoid regulated kinase	71.16	3.42
	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	1.35	1.34
	426067	AA401369	Hs.190721	ESTs	1.01	17.00
55	426088	AF038007	Hs.166196	ATPase, Class I, type 8B, member 1	26.26	47.00
	426215	AW067800	Hs.155223	stanniocalcin 2	1.91	2.90
	426227	U67058	Hs.154299	Human proteinase activated receptor-2 mR	22.40	25.00
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	1.00	1.00
60	426283	NM_003937	Hs.169139	kynureninase (L-kynurenine hydrolase)	91.39	229.00
	426329	AL389951	Hs.271623	nucleoporin 50kD	4.34	4.08
	426427	M86699	Hs.169840	TTK protein kinase	7.02	1.00
	426432	AF001601	Hs.169857	paraoxonase 2	1.16	1.68
	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu	2.59	1.71
	426459	AF151812	Hs.169992	hypothetical 43.2 Kd protein	1.56	1.66
65	426471	M22440	Hs.170009	transforming growth factor, alpha	20.60	25.00
	426496	D31765	Hs.170114	KIAA0061 protein	9.81	22.00
	426501	AA401369	Hs.190721	ESTs	19.23	17.00
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	103.74	41.00
	426536	A1949749	Hs.44441	ESTs	4.65	23.00
70	426572	AB037783	Hs.170623	hypothetical protein FLJ11183	1.00	43.00
	426682	AV660038	Hs.2055	UDP glycosyltransferase 1 family, polype	160.06	8.00
	426691	NM_006201	Hs.171834	PCTAIRE protein kinase 1	1.51	1.35
	426746	J03626	Hs.2057	uridine monophosphate synthetase (arotai	2.13	1.68
	426752	X69490	Hs.172004	titin	0.02	5.14
75	426784	U03749	Hs.172216	chromogranin A (parathyroid secretory pr	1.72	1.71
	426807	AA385315	Hs.156682	ESTs	1.30	1.64
	426812	AF105365	Hs.172613	solute carrier family 12 (potassium/chlo	1.47	1.53
	426814	AF036943	Hs.172619	myelin transcription factor 1-like	1.00	1.00
	426831	BE296216	Hs.172673	S-adenosylhomocysteine hydrolase	1.51	1.25
80	426897	AA401369	Hs.190721	ESTs	141.56	17.00
	426925	NM_001196	Hs.315589	Homo sapiens cDNA: FLJ22373 fis, clone H	32.61	38.00
	426935	NM_000088	Hs.172928	collagen, type I, alpha 1	2.65	3.16
	426964	AA393739	Hs.287416	Homo sapiens cDNA FLJ11439 fis, clone HE	1.97	3.49
	426966	A193134		sclerostin	1.00	1.00
85	426991	AK001536		Homo sapiens cDNA FLJ10674 fis, clone NT	3.39	2.28
	427099	AB032953	Hs.173560	odd Oz/ten-m homolog 2 (Drosophila, mous	4.24	17.00

427239	BE270447	Hs.174070	ubiquitin carrier protein	1.58	1.05
427260	AA653848		gb:ae70b06.s1 Stratagene schizo brain S1	1.34	1.60
427281	AA906147	Hs.102869	ESTs	1.00	66.00
427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
427354	T57896	Hs.191095	ESTs	1.17	1.95
427356	AWD23482	Hs.97849	ESTs	7.31	41.00
427376	AA401533	Hs.19440	ESTs	1.00	57.00
427383	NM_005411	Hs.177582	surfactant, pulmonary-associated protein	0.42	1.32
427427	AF077345	Hs.177936	lectin, superfamily member 1 (cartilage-	1.00	20.00
427441	AA412605	Hs.343879	SPANX family, member C	1.00	1.00
427445	X80818	Hs.178078	glutamate receptor, metabotropic 4	0.97	1.03
427505	AA361562	Hs.178761	26S proteasome-associated pad1 homolog	4.60	4.04
427510	Z47542	Hs.179312	small nuclear RNA activating complex, po	22.00	45.00
427528	AU077143	Hs.179565	minichromosome maintenance deficient (S,	97.45	92.00
427546	AA188763	Hs.36793	hypothetical protein FLJ23188	1.50	3.24
427562	R56424	Hs.26534	ESTs	6.81	40.00
427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	69.91	62.00
427660	A1741320	Hs.114121	Homo sapiens cDNA: FLJ23228 fis, clone C	2.70	49.00
427666	A1791495	Hs.180142	calmodulin-like skin protein	1.37	1.88
427668	AA298760	Hs.180191	hypothetical protein FLJ14904	29.55	67.00
427677	NM_007045	Hs.180295	FGFR1 oncogene partner	3.52	2.63
427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
427711	M31659	Hs.180408	solute carrier family 25 (mitochondrial	15.84	70.00
427719	A1393122	Hs.134726	ESTs	7.03	4.52
427722	AK000123	Hs.180479	hypothetical protein FLJ20116	2.92	1.74
427747	AW411425	Hs.180555	serine/threonine kinase 12	1.76	1.26
427912	AL022310	Hs.181097	tumor necrosis factor (ligand) superfam	9.63	59.00
427961	AW293165	Hs.143134	ESTs	41.97	118.00
428004	AA449563	Hs.151393	glutamate-cysteine ligase, catalytic sub	23.82	1.00
428023	AL038843		Homo sapiens cDNA: FLJ23602 fis, clone L	1.40	1.33
428046	AW812795	Hs.337534	ESTs, Moderately similar to I38022 hypot	96.28	167.00
428093	AW594506	Hs.104830	ESTs	1.25	1.29
428098	AU077258	Hs.182429	protein disulfide isomerase-related prot	1.88	1.60
428129	A1244311	Hs.26512	ESTs	1.00	42.00
428169	A1928984	Hs.182793	golgi phosphoprotein 2	2.76	2.11
428182	BE386042	Hs.293317	ESTs, Weakly similar to GGCI_HUMAN G ANT	1.00	1.00
428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	85.59	181.00
428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	8.57	21.64
428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	7.77	15.90
428434	A1909935	Hs.65551	Homo sapiens, Similar to DNA segment, Ch	0.58	1.43
428450	NM_014791	Hs.184339	KIAA0175 gene product	237.53	204.00
428471	X57348	Hs.184510	stratillin	6.00	4.60
428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	56.54	16.00
428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	3.53	2.15
428505	AL035461	Hs.2281	chromogranin B (secretogranin 1)	1.00	1.00
428532	AF157326	Hs.184786	TBP-interacting protein	1.00	58.00
428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila))-like	1.00	1.00
428698	AA852773	Hs.334938	KIAA1866 protein	187.37	255.00
428728	NM_016625	Hs.191381	hypothetical protein	47.24	80.00
428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.06	1.13
428771	AB028992	Hs.193143	KIAA1069 protein	1.98	92.00
428801	AW277121	Hs.254881	ESTs	1.67	6.15
428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep	1.03	1.27
428839	A1767756	Hs.82302	Homo sapiens cDNA FLJ14814 fis, clone NT	124.17	43.00
428845	AL157579	Hs.153610	KIAA0751 gene product	1.00	1.00
428959	AF100779	Hs.194680	WNT1 inducible signaling pathway protein	15.16	27.00
428969	AF120274	Hs.194689	artemin	1.36	1.24
429038	AL023513	Hs.194766	seizure related gene 6 (mouse)-like	0.97	3.31
429065	A1753247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	6.82	16.47
429164	A1688663	Hs.116586	ESTs	19.08	67.00
429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
429183	AB014604	Hs.197955	KIAA0704 protein	79.72	104.00
429201	X03178	Hs.198246	group-specific component (vitamin D bind	1.00	1.00
429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	1.33	1.09
429220	AW207206		ESTs	1.00	7.00
429228	A1553633	Hs.326447	ESTs	39.47	29.25
429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
429276	AF056085	Hs.198612	G protein-coupled receptor 51	3.70	142.00
429359	W00482	Hs.2399	matrix metalloproteinase 14 (membrane-in	1.30	1.94
429412	NM_006235	Hs.2407	POU domain, class 2, associating factor	94.09	86.00
429413	NM_014058	Hs.201877	DESC1 protein	41.91	10.00
429486	AF155827	Hs.203963	hypothetical protein FLJ10339	12.19	1.00
429504	X99133	Hs.204238	lipocalin 2 (oncogene 24p3)	1.61	1.08
429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
429547	AA401369	Hs.190721	ESTs	1.08	17.00
429551	AW450624	Hs.220931	ESTs	2.89	65.00
429563	BE619413	Hs.2437	eukaryotic translation initiation factor	1.49	1.37
429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	61.86	100.00
429610	AB024937	Hs.211092	LUNX protein; FLUNC (palate lung and nas	1.59	1.69
429612	AF052649	Hs.252587	pituitary tumor-transforming 1	2.78	1.74
429616	A1982722	Hs.120845	ESTs	1.00	1.00
429656	X05608	Hs.211584	neurofilament, light polypeptide (68kD)	1.00	4.00



	429663	M58874	Hs.211587	phospholipase A2, group IVA (cytosolic,	69.95	104.00
	429736	AF125304	Hs.212680	tumor necrosis factor receptor superfamily	1.25	1.21
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	1.00	7.00
5	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory subunit	11.80	1.00
	429918	AW873986	Hs.119383	ESTs	1.00	78.00
	429978	AA249027		ribosomal protein S6	1.98	3.09
	429986	AF092047	Hs.227277	sine oculis homeobox (Drosophila) homolog	1.00	48.00
	430044	AA464510	Hs.152812	ESTs	69.27	59.00
10	430114	AA847744	Hs.99640	ESTs	1.00	1.00
	430134	BE380149	Hs.105223	ESTs, Weakly similar to T33188 hypothetical	1.00	51.00
	430147	R60704	Hs.234434	hair/enhancer-of-split related with YRP	1.10	2.22
	430287	AW182459	Hs.125759	ESTs, Weakly similar to LEU5_HUMAN LEUKE	1.00	127.00
	430294	AI538226	Hs.32976	guanine nucleotide binding protein 4	3.80	1.47
	430300	U60805	Hs.238648	oncostatin M receptor	1.00	35.00
15	430315	NM_004293	Hs.239147	guanine deaminase	92.31	28.00
	430337	M36707	Hs.239600	calmodulin-like 3	1.18	1.08
	430378	Z29572	Hs.2556	tumor necrosis factor receptor superfamily	5.28	66.00
	430388	AA356923	Hs.240770	nuclear cap binding protein subunit 2, 2	16.76	38.00
20	430393	BE185030	Hs.241305	estrogen-responsive B box protein	1.63	1.60
	430439	AL133561		DKFZP434B061 protein	1.00	1.00
	430451	AA836472	Hs.297939	cathepsin B	1.64	2.12
	430454	AW469011	Hs.105535	ESTs	63.35	44.00
	430466	AF052573	Hs.241517	polymerase (DNA directed), theta	2.47	1.91
25	430481	AA479678	Hs.203269	ESTs, Moderately similar to ALU8_HUMAN A	1.00	31.00
	430486	BE062109	Hs.241551	chloride channel, calcium activated, family	12.28	41.00
	430508	AI015435	Hs.104637	ESTs	4.75	7.27
	430533	AA480895	Hs.57749	ESTs, Weakly similar to T17288 hypothetical	1.00	1.00
	430563	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	1.00	1.69
	430677	Z26317	Hs.94560	desmoglein 2	1.72	1.30
30	430678	AA401369	Hs.190721	ESTs	0.90	17.00
	430686	NM_001942	Hs.2633	desmoglein 1	1.00	1.00
	430788	AI742925	Hs.7179	ESTs, Weakly similar to 2004399A chromos	1.62	1.84
	430890	X54232	Hs.2699	glycocalyx 1	1.58	1.40
35	430935	AW072916		zinc finger protein 131 (clone pHZ-10)	90.28	132.00
	430985	AA490232	Hs.27323	ESTs, Weakly similar to 178885 serine/th	0.94	1.26
	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
	431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
	431092	AI332764	Hs.125757	ESTs	13.46	63.00
40	431124	AF284221	Hs.59506	doublesex and mab-3 related transcription	49.43	62.00
	431164	AA493650	Hs.94387	Homo sapiens cDNA: FLJ23494 f1s, clone L	0.44	2.20
	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	182.26	101.00
	431221	AW207837	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye	4.15	13.97
	431277	AA501806	Hs.345824	ESTs	1.00	86.00
45	431322	AW970622		gb:EST382704 MAGE resequences, MAGK Homo	40.55	200.00
	431342	AW971018	Hs.21659	ESTs	1.00	53.00
	431384	BE158000	Hs.285026	gb:MR2-HT0377-150200-202-e03 HT0377 Homo	0.94	1.14
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	1.30	1.25
	431494	AA991355	Hs.298312	hypothetical protein DKFZp434A1315	3.90	26.00
50	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	1.41	1.87
	431548	AJ834273	Hs.97111	novel protein	5.66	15.00
	431630	NM_002204	Hs.265829	integrin, alpha 3 (CD49C, alpha	0.99	1.44
	431745	AW972448	Hs.163425	ESTs	0.99	3.51
	431770	BE221880	Hs.268555	5'-3' exonuclease 2	67.12	91.00
55	431830	Y16645	Hs.271387	small inducible cytokine subfamily A (Cy	3.36	4.71
	431846	BE019924	Hs.271580	uroplakin 1B	4.49	2.51
	431890	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	2.20	3.32
	431934	AB031481	Hs.272214	STG protein	1.01	1.04
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	51.17	46.35
60	432006	AL137382	Hs.272320	Homo sapiens mRNA; cDNA DKFZp434L1226 (f	0.94	1.65
	432023	R43020	Hs.236223	EST	0.94	47.00
	432201	AJ538613	Hs.298241	Transmembrane protease, serine 3	1.10	2.24
	432210	AJ567421	Hs.273330	Homo sapiens, clone IMAGE:3544662, mRNA,	1.42	1.45
	432226	AW182766	Hs.273558	phosphate cytidylyltransferase 1, cholin	1.00	1.00
65	432239	X81334	Hs.2936	matrix metalloproteinase 13 (collagenase	18.67	1.00
	432265	BE382679	Hs.285753	SCG10-like-protein	1.09	1.21
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	40.98	58.00
	432365	AK001106	Hs.274419	hypothetical protein FLJ10244	1.00	214.00
	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 f1s, clone PL	157.34	37.00
70	432376	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
	432407	AA221036		gb:zr03f12.1 Stratagene NT2 neuronal pr	73.71	75.00
	432441	AW292425	Hs.163484	ESTs	56.35	72.00
	432489	AI804855	Hs.207530	ESTs	1.00	24.00
	432543	AA552690	Hs.152423	Homo sapiens cDNA: FLJ21274 f1s, clone C	137.72	98.00
75	432552	AI537170	Hs.173725	ESTs, Weakly similar to ALU8_HUMAN ALU S	1.00	31.00
	432583	AW023524	Hs.162282	potassium channel TASK-4; potassium chan	0.27	35.18
	432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	2.87	6.22
	432625	AI243596	Hs.94830	ESTs, Moderately similar to T03094 A-kin	26.63	56.00
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	1.92	5.29
80	432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	45.13	31.00
	432753	NM_014075	Hs.336938	Homo sapiens PRO0593 mRNA, complete cds	1.00	68.00
	432788	AA521091	Hs.178499	Homo sapiens cDNA: FLJ23117 f1s, clone L	2.69	3.67
	432842	AW674093	Hs.334822	hypothetical protein MGC4485	1.22	1.34
85	432867	AW018936	Hs.233364	ESTs	1.00	1.00
	432917	NM_014125	Hs.241517	PRO0327 protein	10.25	6.62



	432920	U37689	Hs.3128	polymerase (RNA) II (DNA directed) polyp	1.44	1.30
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	164.79	85.64
	433023	AW864793	Hs.87409	thrombospondin 1	20.96	100.00
5	433042	AW193534	Hs.281896	Homo sapiens cDNA FLJ11660 fis, clone HE	1.00	10.00
	433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
	433159	AB035898	Hs.150587	kinesin-like protein 2	13.82	39.00
	433183	AF231338	Hs.222024	transcription factor BMAL2	1.00	69.00
	433258	AA622788	Hs.203613	ESTs, Weakly similar to ALUB_HUMAN III	1.00	1.25
10	433409	AI278802	Hs.25661	ESTs	44.81	117.00
	433437	U20536	Hs.3280	caspase 6, apoptosis-related cysteine pr	70.39	105.00
	433485	AI493076	Hs.201967	aldo-keto reductase family 1, member C2	11.55	2.00
	433537	AI733692	Hs.112488	ESTs	8.66	55.00
	433547	W04978	Hs.303023	beta tubulin 1, class VI	25.16	83.00
	433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
15	433647	AA603367	Hs.222294	ESTs	20.30	49.00
	433658	L03678	Hs.156110	immunoglobulin kappa constant	5.92	10.03
	433800	AI094221	Hs.135150	lung type-I cell membrane-associated gly	2.29	2.22
	433819	AW511097	Hs.112765	ESTs	3.71	8.00
	433862	D86960	Hs.3610	KIAA0205 gene product	62.08	104.00
20	433980	AA137152	Hs.286049	phosphoserine aminotransferase	108.91	47.00
	434088	AF116677	Hs.249270	hypothetical protein PRO1966	1.00	1.00
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	121.27	87.00
	434105	AW952124	Hs.13094	presenilins associated rhomboid-like pro	1.22	1.23
	434217	AW014795	Hs.23349	ESTs	14.11	57.00
25	434340	AI193043	Hs.128685	ESTs, Weakly similar to T17226 hypotheti	2.10	2.56
	434360	AA401369	Hs.190721	ESTs	40.98	17.00
	434414	AI798376		gbtr34b07.x1 NCI_CGAP_Ov23 Homo sapiens	1.48	1.56
	434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	64.00
	434467	BE552368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	54.91	85.00
30	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	2.46	2.00
	434627	AI221894	Hs.39311	ESTs	1.00	1.00
	434699	AA643687	Hs.149425	Homo sapiens cDNA FLJ11980 fis, clone HE	1.00	23.00
	434769	AA648884	Hs.134278	Homo sapiens cDNA FLJ12676 fis, clone NT	7.08	56.00
	434792	AA649253	Hs.132458	ESTs	8.52	44.00
35	434808	AF155108	Hs.256150	Homo sapiens, Similar to RIKEN cDNA 2810	11.33	1.00
	434828	D90070	Hs.96	phorbol-12-myristate-13-acetate-induced	1.00	1.00
	434876	AF160477	Hs.61460	Ig superfamily receptor LNIR	1.25	1.29
	434891	AA814309	Hs.123583	ESTs	1.00	6.00
	434928	AW015595	Hs.4267	Homo sapiens clones 24714 and 24715 mRNA	1.00	1.00
40	435013	H91923	Hs.110024	Target CAT	1.26	1.10
	435066	BE261760	Hs.4747	dyskeratosis congenita 1, dyskerin	1.69	1.37
	435087	AW975241	Hs.23567	ESTs	1.00	1.00
	435099	AC004770	Hs.4756	flap structure-specific endonuclease 1	2.90	1.93
45	435159	AA668879	Hs.116649	ESTs	1.00	1.00
	435205	X54136	Hs.181125	immunoglobulin lambda locus	1.02	1.46
	435232	NM_001262	Hs.4854	cyclin-dependent kinase inhibitor 2C (p1	2.04	2.70
	435304	H10709	Hs.269524	ESTs	27.58	139.00
	435313	AI769400	Hs.189729	ESTs	1.00	14.00
	435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
50	435509	AI458679	Hs.181915	ESTs	1.00	1.00
	435525	AI831297	Hs.123310	ESTs	1.00	56.00
	435532	AW291488	Hs.117305	Homo sapiens, clone IMAGE:3682908, mRNA	1.00	2.00
	435550	AI224458	Hs.324507	H.sapiens polyA site DNA	3.42	3.92
	435602	AF217515	Hs.283532	uncharacterized bone marrow protein BM03	3.95	1.80
55	435766	R11673	Hs.186498	ESTs	1.00	28.00
	435793	AB037734	Hs.4993	KIAA1313 protein	23.68	42.00
	436069	AI056879	Hs.263209	ESTs	1.00	58.00
	436170	AW450381	Hs.14529	ESTs	1.00	18.00
	436211	AK001581	Hs.334828	hypothetical protein FLJ10719; KIAA1794	5.84	22.00
60	436213	AA325512	Hs.71472	hypothetical protein FLJ10774; KIAA1709	1.42	1.27
	436217	T53925	Hs.107	fibrinogen-like 1	57.97	31.00
	436238	AK002163	Hs.301724	hypothetical protein FLJ11301	2.51	1.71
	436251	BE515065	Hs.296585	nucleolar protein (KKE/D repeat)	2.33	1.64
	436291	BE568452	Hs.344037	protein regulator of cytokinesis 1	108.99	52.00
65	436302	AL355841	Hs.99330	hypothetical protein FLJ23588	0.75	2.81
	436396	AW992292	Hs.152213	wingless-type MMTV Integration site fami	60.01	1.00
	436414	BE284633	Hs.143638	WD repeat domain 4	2.50	2.19
	436419	AI948626	Hs.171356	ESTs	0.95	1.33
	436443	AW138211	Hs.128746	ESTs	1.12	9.26
70	436474	AJ270693	Hs.199887	ESTs	1.00	1.00
	436481	AA379597	Hs.5189	HSPC150 protein similar to ubiquitin-con	3.28	1.58
	436486	AA742221	Hs.120633	ESTs	1.00	19.00
	436511	AA721252	Hs.291502	ESTs	16.76	14.00
	436553	X57809	Hs.181125	immunoglobulin lambda locus	1.08	1.74
75	436557	W15573	Hs.5027	ESTs, Weakly similar to A47582 B-cell gr	19.20	9.75
	436608	AA626980		down syndrome critical region protein DS	33.92	25.00
	436667	AW025183	Hs.127680	ESTs	0.89	1.19
	436771	AW975687	Hs.292979	ESTs	1.00	10.00
	436839	AA401369	Hs.190721	ESTs	1.00	17.00
80	436887	AW953157	Hs.193235	hypothetical protein DKFZp547D155	1.06	1.15
	436944	AW268614	Hs.5840	ESTs	1.00	1.00
	436961	AW375974	Hs.156704	ESTs	25.13	25.00
	436972	AA284679	Hs.25640	claudin 3	1.59	1.46
	437016	AU076916	Hs.5398	guanine monophosphate synthetase	2.35	1.78
85	437044	AL035864	Hs.69517	cDNA for differentially expressed CO16 g	1.34	1.13

	437181	AI306615	Hs.125343	ESTs, Weakly similar to KIAA0758 protein	1.00	17.00
	437204	AL110216	Hs.22826	ESTs, Weakly similar to I55214 salivary	40.55	82.00
	437205	AL110232	Hs.279243	Homo sapiens mRNA; cDNA DKFZp564D2071 (f	1.00	112.00
5	437259	AI377755	Hs.120695	ESTs	1.00	205.00
	437270	R18087	Hs.323769	cisplatin resistance related protein CRR	1.56	1.54
	437271	AL137445	Hs.28846	Homo sapiens mRNA; cDNA DKFZp566O134 (fr	113.25	125.00
	437370	AL359567	Hs.161962	Homo sapiens mRNA; cDNA DKFZp547D023 (fr	1.82	4.57
	437390	AI125859	Hs.112607	ESTs	1.35	1.75
10	437412	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (fr	3.58	3.20
	437435	AI306152	Hs.27027	hypothetical protein DKFZp762H1311	3.03	1.08
	437444	HA6008	Hs.31518	ESTs	1.00	39.00
	437568	AI954795	Hs.156135	ESTs	1.00	19.00
	437623	D63880	Hs.5719	chromosome condensation-related SMC-asso	1.95	1.57
	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti	1.00	3.00
15	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	1.00	45.00
	437840	AA884836	Hs.292014	ESTs	1.07	1.78
	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.s]	1.68	3.26
	437879	BE262082	Hs.5894	hypothetical protein FLJ10305	1.87	2.52
20	437915	AI637993	Hs.202312	Homo sapiens clone N11 NTERA2D1 teratoca	74.05	35.00
	437916	BE566249	Hs.20999	hypothetical protein FLJ23142	23.15	89.00
	437937	AI917222	Hs.121655	ESTs	1.00	1.00
	437942	AI888256	Hs.307526	ESTs	12.28	31.00
	438091	AW373062		nuclear receptor subfamily 1, group I, m	1.53	10.85
25	438113	AI467908	Hs.8882	ESTs	1.80	2.39
	438119	AW963217	Hs.203961	ESTs, Moderately similar to AF116721 89	22.67	36.90
	438274	AI918906	Hs.55080	ESTs	1.00	1.00
	438378	AW970529	Hs.86434	hypothetical protein FLJ21816	38.92	38.00
	438403	AA806607	Hs.292206	ESTs	1.00	1.00
30	438494	AA909678	Hs.130183	ESTs	2.05	80.00
	438546	AW297204	Hs.125811	ESTs	1.00	131.00
	438552	AI245820	Hs.6314	type I transmembrane receptor (seizure-r	1.43	1.45
	438702	AI879064	Hs.54618	ESTs	1.00	34.00
	438724	AW612553	Hs.114670	Human DNA sequence from clone RP11-16L21	1.33	1.10
35	438746	AI885815	Hs.184727	Human melanoma-associated antigen p97 (m	2.42	1.59
	438779	NM_003787	Hs.6414	nuclear protein 4	1.00	18.00
	438821	AA826425	Hs.192375	ESTs	2.03	2.57
	438885	AI886558	Hs.184987	ESTs	6.42	88.00
	438898	AA401369	Hs.190721	ESTs	22.41	17.00
40	438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regi	1.00	1.00
	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9	2.20	1.88
	439000	AW979121		gh:EST391231 MAGE resequences, MAGP Homo	2.78	4.81
	439023	AA745978	Hs.28273	ESTs	1.17	1.31
	439024	R96696	Hs.35598	ESTs	1.00	28.00
45	439128	AI943371	Hs.153089	ESTs	1.00	67.00
	439146	AW138909	Hs.156110	immunoglobulin kappa constant	1.38	1.41
	439223	AW238299	Hs.250618	UL16 binding protein 2	1.93	1.64
	439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
50	439318	AW837046	Hs.6527	G protein-coupled receptor 56	2.00	2.20
	439343	AF086161	Hs.114611	hypothetical protein FLJ11808	6.10	7.37
	439394	AA401369	Hs.190721	ESTs	3.39	17.00
	439410	AA632012	Hs.188746	ESTs	1.83	3.07
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00
	439452	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	18.76	122.00
55	439453	BE264974	Hs.6566	thyroid hormone receptor interactor 13	2.78	1.58
	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	1.22	1.44
	439492	AF086310	Hs.103159	ESTs	7.46	39.00
	439523	W72348	Hs.185029	ESTs	1.00	1.19
	439592	AF086413	Hs.58399	ESTs	1.00	1.00
60	439606	W79123	Hs.58581	G protein-coupled receptor 87	33.61	1.00
	439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
	439702	AW085525	Hs.134182	ESTs	4.30	10.00
	439706	AW872527	Hs.59761	ESTs, Weakly similar to DAPI1_HUMAN DEATH	86.55	11.00
	439738	BE246502	Hs.9598	sera domain, immunoglobulin domain (Ig),	2.36	1.88
65	439750	AL359053	Hs.57664	Homo sapiens mRNA full length insert cDN	2.02	6.08
	439759	AL359055	Hs.57709	Homo sapiens mRNA full length insert cDN	1.00	21.00
	439780	AL109688		gh:Homo sapiens mRNA full length insert	7.27	25.00
	439840	AW449211	Hs.105445	GDNF family receptor alpha 1	1.00	1.00
	439926	AW014875	Hs.137007	ESTs	32.58	71.00
70	439953	AW247529	Hs.6793	platelet-activating factor acetylhydrola	21.28	9.55
	439979	AW600291	Hs.6823	hypothetical protein FLJ10430	68.83	61.00
	440006	AK000517	Hs.6844	hypothetical protein FLJ20510	1.83	4.02
	440028	AW473675	Hs.125843	ESTs, Weakly similar to T17227 hypotheti	1.42	2.54
	440106	AA864968	Hs.127699	KIAA1603 protein	1.00	54.00
75	440138	AB033023	Hs.318127	hypothetical protein FLJ10201	24.18	52.00
	440273	AI805392	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	3.21	4.72
	440289	AW450991	Hs.192071	ESTs	38.63	113.00
	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
	440492	R39127	Hs.21433	hypothetical protein DKFZp547J036	2.35	3.62
80	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	10.84	57.00
	440659	AF134160	Hs.7327	claudin 1	3.18	2.37
	440704	M69241	Hs.162	insulin-like growth factor binding prote	2.89	2.09
	440943	AW082298	Hs.146161	hypothetical protein MGC2408	2.02	1.41
	440994	AI160011	Hs.272068	ESTs	1.29	1.14
85	441020	AA401369	Hs.190721	ESTs	142.99	17.00
	441031	AI110694	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00

	441128	AA570256	ESTs, Weakly similar to T23273 hypotheti	4.13	3.50
	441290	W27501	Hs.89605 cholinergic receptor, nicotinic, alpha p	1.00	1.00
	441362	BE614410	Hs.23044 RAD51 (S. cerevisiae) homolog (E coli Re	130.23	43.00
5	441377	BE218239	Hs.202656 ESTs	22.03	1.00
	441390	AI692560	Hs.131175 ESTs	3.65	7.70
	441497	R51064	Hs.23172 ESTs	1.00	1.00
	441525	AW241867	Hs.127728 ESTs	1.63	1.42
	441553	AA281219	Hs.121296 ESTs	1.89	1.57
	441607	NM_005010	Hs.7912 neuronal cell adhesion molecule	1.47	2.11
10	441633	AW958544	Hs.112242 normal mucosa of esophagus specific 1	216.22	363.00
	441636	AA081846	Hs.7921 Homo sapiens mRNA; cDNA DKFZp566E183 (fr	2.31	2.05
	441737	X79449	Hs.7957 adenosine deaminase, RNA-specific	1.30	1.49
	441790	AA401369	Hs.190721 ESTs	44.15	17.00
	441801	AW242799	Hs.86366 ESTs	1.00	1.00
15	441919	AI553802	Hs.128121 ESTs	1.00	122.00
	441937	R41782	Hs.22279 ESTs	0.86	1.37
	441954	AI744935	Hs.8047 Fanconi anemia, complementation group G	1.48	1.39
	442025	AW887434	Hs.11810 CDA11 protein	1.00	46.00
20	442029	AW956698	Hs.14456 neural precursor cell expressed, develop	9.92	45.00
	442072	AJ740832	Hs.12311 Homo sapiens clone 23570 mRNA sequence	25.05	77.00
	442108	AW452649	Hs.166314 ESTs	3.61	3.14
	442117	AW664964	Hs.128899 ESTs	3.00	5.49
	442137	AA977235	Hs.128830 ESTs, Weakly similar to Z192_HUMAN ZINC	1.00	1.00
25	442159	AW163390	Hs.278554 heterochromatin-like protein 1	1.92	1.66
	442179	AA983842	Hs.333555 chromosome 2 open reading frame 2	27.22	50.00
	442328	AI952430	Hs.150614 ESTs, Weakly similar to ALU4_HUMAN ALU S	5.00	3.42
	442432	BE093589	Hs.38178 hypothetical protein FLJ23458	181.59	76.00
	442530	AI580830	Hs.176508 Homo sapiens cDNA FLJ14712 fis, clone NT	10.59	144.00
30	442547	AA306897	Hs.217484 ESTs, Weakly similar to ALU1_HUMAN ALU S	109.23	98.00
	442556	AL137761	Hs.8379 Homo sapiens mRNA; cDNA DKFZp586L2424 (f	1.00	53.00
	442619	AA447492	Hs.20183 ESTs, Weakly similar to AF164793 1 prote	29.02	50.00
	442710	AI015631	Hs.23210 ESTs	1.00	19.00
	442717	R88362	Hs.180591 ESTs, Weakly similar to T23976 hypotheti	1.00	5.00
35	442875	BE623003	Hs.23625 Homo sapiens clone TCCCTA00142 mRNA sequ	22.85	50.00
	442914	AW188551	Hs.99519 hypothetical protein FLJ14007	25.33	82.00
	442932	AA457211	Hs.8858 bromodomain adjacent to zinc finger doma	3.18	4.41
	442942	AW167087	Hs.131562 ESTs	8.45	64.00
	443068	AI188710	ESTs	1.00	27.00
40	443204	AW205876	Hs.29643 Homo sapiens cDNA FLJ13103 fis, clone NT	1.00	24.00
	443211	AI128388	Hs.143655 ESTs	12.42	2.00
	443247	BE614387	Hs.333893 c-Myc target JPO1	128.84	96.00
	443324	R44013	Hs.164225 ESTs	0.02	4.59
	443383	AI792453	Hs.168507 ESTs	1.00	47.00
45	443400	R28424	Hs.250648 ESTs	18.52	61.00
	443426	AF098158	Hs.9329 chromosome 20 open reading frame 1	4.02	1.75
	443572	AA025610	Hs.9605 cleavage and polyadenylation specific fa	2.98	2.57
	443575	AI078022	Hs.269636 ESTs, Weakly similar to ALU1_HUMAN ALU S	1.00	29.00
	443614	AV655386	Hs.7645 fibrinogen, B beta polypeptide	1.00	16.00
	443633	AL031290	Hs.9654 similar to pregnancy-associated plasma p	1.00	39.00
50	443648	AI085377	Hs.143810 ESTs	39.81	70.00
	443715	AI583187	Hs.9700 cyclin E1	48.74	7.00
	443723	AI144442	Hs.157144 syntaxin 6	1.29	1.30
	443802	AW504924	Hs.9805 KIAA1291 protein	1.75	1.61
	443859	NM_013409	Hs.9914 follistatin	1.35	1.13
55	443892	AA401369	Hs.190721 ESTs	1.00	17.00
	443947	W24187	gbzb4709.1 Soares_fetal_lung_NbHL19W	1.33	1.64
	443991	NM_002250	Hs.10082 potassium intermediate/small conductance	5.71	6.87
	444006	BE395085	Hs.10088 type I transmembrane protein Fn14	1.47	1.92
60	444009	AI380792	Hs.135104 ESTs	1.00	77.00
	444017	U04840	Hs.214 neuro-oncological ventral antigen 1	1.00	1.00
	444127	N63620	Hs.13281 ESTs	1.00	29.00
	444129	AW294292	Hs.256212 ESTs	1.00	1.00
	444279	U62432	Hs.89605 cholinergic receptor, nicotinic, alpha p	0.60	7.80
65	444371	BE540274	Hs.239 forkhead box M1	2.91	1.14
	444378	R41339	Hs.12569 ESTs	1.00	1.00
	444381	BE387335	Hs.283713 ESTs, Weakly similar to S64054 hypotheti	469.00	556.00
	444461	R53734	Hs.25978 ESTs, Weakly similar to 2109260A B cell	12.88	105.00
	444471	AB020684	Hs.11217 KIAA0877 protein	24.91	90.00
	444489	AI151010	Hs.157774 ESTs	1.00	111.00
70	444619	BE538082	Hs.8172 ESTs, Moderately similar to A46010 X-in	1.00	70.00
	444665	BE613126	Hs.47783 B aggressive lymphoma gene	30.56	139.00
	444707	AI188613	Hs.41690 desmocollin 3	1.00	1.00
	444735	BE019923	Hs.243122 hypothetical protein FLJ13057 similar to	77.02	90.00
75	444781	NM_014400	Hs.11950 GPI-anchored metastasis-associated prote	1.57	1.31
	444783	AK001468	Hs.62180 anillin (Drosophila Scraps homolog), act	77.55	2.00
	445236	AK001676	Hs.12457 hypothetical protein FLJ10814	1.00	27.00
	445258	AI635931	Hs.147613 ESTs	1.00	73.00
	445413	AA151342	Hs.12677 CGI-147 protein	28.14	50.00
	445417	AK001058	Hs.12680 Homo sapiens cDNA FLJ10196 fis, clone HE	1.81	2.62
80	445443	AV653838	Hs.322971 ESTs	1.00	1.00
	445462	AA378776	Hs.288649 hypothetical protein MGC3077	2.09	1.70
	445517	AF208855	Hs.12830 hypothetical protein	1.87	70.00
	445537	AJ245671	Hs.12844 EGF-like domain, multiple 6	1.71	2.72
	445580	AF167572	Hs.12912 skb1 (S. pombe) homolog	1.52	1.34
85	445654	X91247	Hs.13046 thioredoxin reductase 1	1.51	1.52

	445669	AI570830	Hs.174870	ESTs	10.95	11.45
	445818	BE045321	Hs.136017	ESTs	1.00	1.00
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 1-I	49.42	54.00
5	445885	AI734009	Hs.127699	KIAA1603 protein	1.00	132.00
	445898	AF070623	Hs.13423	Homo sapiens clone 24468 mRNA sequence	1.00	1.00
	445903	AI347487	Hs.132781	class I cytokine receptor	1.00	36.00
	445932	BE046441	Hs.333555	Homo sapiens clone 24859 mRNA sequence	2.41	2.88
	445982	BE410233	Hs.13501	pescadillo (zebrafish) homolog 1, contai	1.60	1.35
10	446078	AI339982	Hs.156061	ESTs	1.00	42.00
	446102	AW168067	Hs.317694	ESTs	1.00	1.00
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	1.70	1.53
	446269	AW263155	Hs.14559	hypothetical protein FLJ10540	73.01	48.00
	446292	AF081497	Hs.279582	Rh type C glycoprotein	1.55	1.28
	446293	AI420213	Hs.149722	ESTs	1.00	2.00
15	446423	AW139555	Hs.150120	ESTs	1.10	4.19
	446428	AW082270	Hs.12496	ESTs, Weakly similar to ALU4_HUMAN ALU S	0.53	3.26
	446432	AI377320	Hs.150059	ESTs	1.00	5.00
	446528	AU076540	Hs.15243	nucleolar protein 1 (120kD)	1.36	1.31
	446574	AI310135	Hs.335933	ESTs	3.89	72.00
20	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	32.03	20.23
	446636	AC002563	Hs.15787	citron (rho-interacting, serine/threonin	4.19	5.07
	446783	AW138343	Hs.141867	ESTs	2.82	9.47
	446839	BE091926	Hs.16244	mitotic spindle coiled-coil related prot	110.28	28.00
	446849	AU076617	Hs.16251	cleavage and polyadenylation specific fa	3.26	2.94
25	446856	AI814373	Hs.164175	ESTs	6.38	11.30
	446872	X97058	Hs.16362	pyrimidinergic receptor P2Y, G-protein c	1.98	2.03
	446880	AI811807	Hs.108648	Homo sapiens cDNA FLJ14934 fis, clone PL	94.90	113.00
	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	1.67	3.90
	446989	AK001898	Hs.16740	hypothetical protein FLJ11036	2.82	3.12
30	447022	AW291223	Hs.157573	ESTs	1.00	170.00
	447033	AI357412	Hs.157601	ESTs	7.15	107.00
	447078	AW885727	Hs.9914	ESTs	47.24	24.00
	447081	Y13696	Hs.17287	potassium inwardly-rectifying channel, s	0.12	17.88
35	447131	NM_004585	Hs.17466	retinoic acid receptor responder (lazarus	0.97	1.48
	447149	BE299857	Hs.326	TAR (HIV) RNA-binding protein 2	1.24	1.26
	447153	AA805202	Hs.315562	ESTs	1.00	54.00
	447164	AF026941	Hs.17518	Homo sapiens cigs mRNA, partial sequence	1.00	67.00
	447178	AW594541	Hs.192417	ESTs	3.42	50.00
40	447250	AI878909	Hs.17883	protein phosphatase 1G (formerly 2C), ma	1.60	1.52
	447289	AW247017	Hs.36978	melanoma antigen, family A, 3	1.00	1.00
	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m	146.62	51.00
	447350	AJ375572	Hs.172634	ESTs	1.00	12.00
	447377	N27687	Hs.334334	transcription factor AP-2 alpha (activat	2.55	63.00
45	447415	AW937335	Hs.28149	ESTs, Weakly similar to KF38_HUMAN KINES	0.91	1.13
	447425	AJ963747	Hs.18573	acylphosphatase 1, erythrocyte (common)	1.00	35.00
	447519	U46258	Hs.339665	ESTs	59.89	49.00
	447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
50	447534	AA401369	Hs.190721	ESTs	1.00	17.00
	447636	Y10043		high-mobility group (nonhistone chromoso	1.41	1.11
	447688	N87079	Hs.19236	Target CAT	1.00	39.00
	447733	AF157482	Hs.19400	MAD2 (mitotic arrest deficient, yeast, h	1.17	1.12
	447769	AW873704	Hs.320831	Homo sapiens cDNA FLJ14597 fis, clone NT	6.47	5.95
55	447802	AW593432	Hs.161455	ESTs	0.73	2.34
	447850	AB018298	Hs.19822	SEC24 (S. cerevisiae) related gene famil	86.45	116.00
	447924	AI817226	Hs.313413	ESTs, Weakly similar to T23110 hypotheti	1.00	1.00
	447973	AB011169	Hs.20141	similar to S. cerevisiae SSNA4	3.50	4.27
	448030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	4.13	142.00
60	448105	AI538613	Hs.298241	Transmembrane protease, serine 3	1.15	2.24
	448243	AW369771	Hs.52620	integrin, beta 8	15.84	1.00
	448278	W07369	Hs.11782	ESTs	0.97	1.90
	448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	1.00	1.00
	448296	BE622756	Hs.10949	Homo sapiens cDNA FLJ14162 fis, clone NT	2.42	2.17
65	448357	BE274396	Hs.108923	RAB38, member RAS oncogene family	1.44	1.08
	448390	AL035414	Hs.21068	hypothetical protein	1.00	43.00
	448469	AW504732	Hs.21275	hypothetical protein FLJ11011	2.63	2.49
	448569	BE382657	Hs.21486	signal transducer and activator of trans	1.84	2.53
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.29	46.00
70	448672	AI955511	Hs.225106	ESTs	1.00	21.00
	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte	1.82	1.08
	448741	BE614567	Hs.19574	hypothetical protein MGC5469	2.48	1.92
	448757	AI366784	Hs.48820	TATA box binding protein (TBP)-associate	23.53	20.00
	448775	AB025237	Hs.388	nucleoside diphosphate linked moi	2.34	1.97
75	448826	AI580252	Hs.293246	ESTs, Weakly similar to putative p150 [H	74.07	62.67
	448830	AL031658	Hs.22181	hypothetical protein dJ310013.3	1.37	1.31
	448844	AI581519	Hs.177164	ESTs	1.00	31.00
	448988	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	1.84	1.95
	448993	AI471630		KIAA0144 gene product	1.63	1.49
80	449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
	449029	N28989	Hs.22891	solute carrier family 7 (cationic amino	1.97	2.26
	449040	AF040704	Hs.149443	putative tumor suppressor	0.97	1.56
	449048	Z45051	Hs.22820	similar to S68401 (cattle) glucose induc	27.13	90.00
	449053	AI625777	Hs.344766	ESTs	8.33	44.00
85	449054	AF148848	Hs.22934	myoneurin	73.85	104.00
	449101	AA205847	Hs.23016	G protein-coupled receptor	2.68	27.00

	449167	T05095	Hs.19597	KIAA1694 protein	1.61	2.36
	449207	AL044222	Hs.23255	nucleoporin 155kD	2.36	1.66
	449228	AJ403107	Hs.148590	protein related with psoriasis	1.15	1.15
5	449230	BE613348	Hs.211579	metanoma cell adhesion molecule	206.65	151.00
	449305	AI638293		gb:109b07.x1 NCL_CGAP_GC6 Homo sapiens	17.28	45.00
	449318	AW235021	Hs.78531	Homo sapiens, Similar to RIKEN cDNA 5730	26.39	35.00
	449448	D60730	Hs.57471	ESTs	1.00	1.00
	449467	AW205006	Hs.197042	ESTs	1.00	1.00
10	449523	NM_000579	Hs.54443	chemokine (C-C motif) receptor 5	56.80	216.86
	449722	BE280074	Hs.23960	cyclin B1	150.03	1.00
	449976	H06350	Hs.135056	Human DNA sequence from clone RP5-850E9	2.16	2.85
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitter)	1.17	1.45
	450098	W27249	Hs.8109	hypothetical protein FLJ121080	1.79	2.38
	450101	AV649989	Hs.24385	Human hbc547 mRNA sequence	1.00	69.00
15	450149	AW969781	Hs.132863	Zic family member 2 (odd-paired Drosophila)	1.00	1.00
	450193	AI916071	Hs.15607	Homo sapiens Fanconi anemia complemental	29.85	34.00
	450221	AA328102	Hs.24541	cytoskeleton associated protein 2	1.00	1.00
	450372	BE218107	Hs.202436	ESTs	1.00	1.00
20	450375	AA009647	Hs.8850	a disintegrin and metalloproteinase domain	51.26	93.00
	450447	AF212223	Hs.25010	hypothetical protein P15-2	123.20	181.00
	450568	AL050078	Hs.25159	Homo sapiens cDNA FLJ10784 fis, clone NT	1.00	19.00
	450589	AI701505	Hs.202526	ESTs	1.00	23.00
	450684	AA872605	Hs.25333	Interleukin 1 receptor, type II	1.00	100.00
	450701	H39960	Hs.288467	Homo sapiens cDNA FLJ12280 fis, clone MA	1.89	1.55
25	450705	U90304	Hs.25351	iroquois homeobox protein 2A (IRX-2A) (	1.00	45.00
	450832	AA401369	Hs.190721	ESTs	25.17	17.00
	450937	R49131	Hs.26267	ATP-dependant Interferon response protein	90.92	90.00
	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	3.33	1.70
30	451105	AI761324		gb:166b011.x1 NCL_CGAP_Co16 Homo sapiens	15.02	124.00
	451110	AI955040	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	143.00
	451253	H48299	Hs.26126	claudin 10	3.02	2.28
	451291	R39288	Hs.6702	ESTs	1.00	1.00
	451320	AW498974		diacylglycerol kinase, zeta (104kD)	2.92	18.00
35	451380	H09280	Hs.13234	ESTs	6.90	6.67
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant)	35.75	72.00
	451437	H24143	Hs.31945	hypothetical protein FLJ11071	1.00	69.00
	451452	AK000367	Hs.26434	hypothetical protein FLJ20360	1.83	2.10
	451524	AK001466	Hs.26516	hypothetical protein FLJ10604	1.13	1.07
40	451541	BE279383	Hs.26557	plakophilin 3	1.88	1.33
	451592	AI805416	Hs.213897	ESTs	1.00	1.00
	451635	AA018899	Hs.127179	cryptic gene	1.52	1.92
	451743	AA401369	Hs.190721	ESTs	4.95	17.00
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	13.55	31.00
45	451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
	451871	AI821005	Hs.118599	ESTs	1.81	2.53
	451952	AL120173	Hs.301663	ESTs	1.00	22.00
	452012	AA307703	Hs.279766	kinasin family member 4A	3.43	2.26
	452046	AB018345	Hs.27657	KIAA0802 protein	56.59	19.00
50	452194	AI694413	Hs.332549	olfactory receptor, family 2, subfamily	1.67	4.09
	452206	AW340281	Hs.33074	Homo sapiens, clone IMAGE:3606519, mRNA,	9.31	53.00
	452240	AA401369	Hs.190721	ESTs	13.42	17.00
	452256	AK000933	Hs.28681	Homo sapiens cDNA FLJ10071 fis, clone HE	39.03	94.00
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fis, clone PL	153.01	340.00
55	452291	AF015592	Hs.28853	CDC7 (cell division cycle 7, S. cerevisiae)	1.95	23.00
	452295	BE379536	Hs.28856	programmed cell death 10	42.33	61.00
	452304	AA025386	Hs.61311	ESTs, Weakly similar to S10590 cysteine	1.17	2.14
	452340	NM_002202	Hs.605	ISL1 transcription factor, LIM/homeodomain	1.00	13.00
	452349	AB028944	Hs.29189	ATPase, Class VI, type 11A	1.09	1.42
60	452367	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	54.49	53.00
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced protein	1.00	32.00
	452410	AL133619		Homo sapiens mRNA; cDNA DKFZp434E2321 (f	1.26	1.99
	452461	N78223	Hs.108106	transcription factor	24.47	35.00
	452571	W31518	Hs.34665	ESTs	54.61	102.00
65	452613	AA461599	Hs.23459	ESTs	1.39	1.32
	452699	AW295390	Hs.213062	ESTs	1.00	26.00
	452705	H49805	Hs.246005	ESTs	1.00	1.00
	452747	AF160477	Hs.61460	lg superfamily receptor LNIR	112.87	1.29
	452787	AW294022	Hs.222707	KIAA1718 protein	1.00	1.00
70	452795	AW392555	Hs.18878	hypothetical protein FLJ21620	1.00	1.00
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	7.91	75.00
	452833	BE559681	Hs.30736	KIAA0124 protein	3.16	1.92
	452838	U65011	Hs.30743	preferentially expressed antigen in melan	174.35	1.00
	452862	AA401369	Hs.190721	ESTs	98.26	17.00
75	452865	AW173720	Hs.345805	ESTs, Weakly similar to A47582 B-cell gr	1.55	1.00
	452934	AA581322	Hs.4213	hypothetical protein MGC16207	1.73	1.19
	452946	X95425	Hs.31092	EphA5	1.00	1.00
	452976	R44214	Hs.101189	ESTs	1.58	1.98
	453028	AB006532	Hs.31442	RacQ protein-like 4	1.80	1.60
80	453095	AW295660	Hs.252756	ESTs	0.77	1.50
	453102	NM_007197	Hs.31664	frizzled (Drosophila) homolog 10	1.00	1.00
	453103	AI301052	Hs.153444	ESTs	1.00	1.00
	453120	AA292891	Hs.31773	pregnancy-induced growth inhibitor	1.23	1.20
	453153	N53893	Hs.24360	ESTs	1.00	83.00
	453160	AI263307	Hs.239884	H2B histone family, member L	1.00	30.00
85	453197	AI916269	Hs.109057	ESTs, Weakly similar to ALU5_HUMAN ALU S	1.00	134.00

5	453210	AL133161	Hs.32360	hypothetical protein FLJ10867	1.69	1.93		
	453240	AI969564	Hs.166254	hypothetical protein DKFZp566i133	1.00	1.00		
	453317	NM_002277	Hs.41696	keratin, hair, acidic, 1	1.19	1.27		
	453323	AF034102	Hs.32951	solute carrier family 29 (nucleoside tra	4.90	4.11		
	453331	AI240665	Hs.8860	ESTs	199.42	340.00		
	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00		
	453431	AF094754	Hs.32973	glycine receptor, beta	1.00	1.00		
	453439	AI572438	Hs.32976	guanine nucleotide binding protein 4	3.44	5.17		
	453459	BE047032	Hs.257789	ESTs	2.84	5.58		
10	453563	AW606906.comp		Hs.181163	hypothetical protein MGC5629	4.58	90.00	
	453633	AA357001	Hs.34045	hypothetical protein FLJ20764	1.74	1.60		
	453775	NM_002918	Hs.35120	replication factor C (activator 1) 4 (37	19.49	1.00		
	453830	AA534296	Hs.20953	ESTs	24.92	25.00		
	453857	AL080235	Hs.35861	DKFZP586E1621 protein	167.59	66.00		
15	453867	AI929363	Hs.33032	hypothetical protein DKFZp434N185	1.00	39.00		
	453883	AI638516	Hs.347524	cofactor required for Sp1 transcriptiona	1.97	1.58		
	453884	AA355925	Hs.36232	KIAA0186 gene product	63.89	20.00		
	453900	AW003582	Hs.226414	ESTs, Weakly similar to ALU8_HUMAN ALU S	20.41	16.00		
	453922	AF053306	Hs.36708	budding uninhibited by benzimidazoles 1	7.09	22.00		
20	453941	U39817	Hs.36820	Bloom syndrome	29.75	19.00		
	453964	AI861486	Hs.12744	ESTs	1.00	1.00		
	453968	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351295, mRNA	2.06	1.81		
	453976	BE463830	Hs.163714	ESTs	3.02	131.00		
25	454024	AA993527	Hs.293907	hypothetical protein FLJ23403	1.00	131.00		
	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member	1.23	1.02		
	454042	T19228	Hs.172572	hypothetical protein FLJ20093	30.63	171.00		
	454059	NM_003154	Hs.37048	stathmin	1.00	1.00		
	454066	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid	1.01	1.45		
30	454098	W27953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11		
	454241	BE144666		gb:CM2-HT0176-041099-017-c02 HT0176 Homo	6.33	5.04		
	454417	AI244459	Hs.110826	trinucleotide repeat containing 9	4.30	7.82		
	454439	AW819152	Hs.154320	DKFZP566O1646 protein	1.00	1.00		
	455175	AW993247		gb:RC2-BN0033-180200-014-h09 BN0033 Homo	13.75	103.00		
35	455601	AI368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00		
	456237	AA203682		gb:zx52e07.r1 Soares_fetal_liver_spleen_	1.00	1.00		
	456321	NM_001327	Hs.87225	cancer/testis antigen	1.14	1.10		
	456475	NM_000144	Hs.95998	Friedreich ataxia	1.00	48.00		
	456508	AA502764	Hs.123469	ESTs, Weakly similar to AF208855 1 BM-01	162.25	189.00		
40	456534	X91195	Hs.100623	phospholipase C, beta 3, neighbor pseudo	2.12	1.80		
	456736	AW248217	Hs.1619	achaete-scute complex (Drosophila) homol	1.15	1.94		
	456759	BE259150	Hs.127792	delta (Drosophila)-like 3	1.00	1.00		
	456990	NM_004504	Hs.171545	HIV-1 Rev binding protein	16.42	84.00		
	457200	U33749	Hs.197764	thyroid transcription factor 1	0.57	1.76		
45	457234	AW968360	Hs.14355	Homo sapiens cDNA FLJ13207 fis, clone NT	2.71	4.15		
	457465	AW301344	Hs.122908	DNA replication factor	46.37	47.00		
	457489	AI693815	Hs.127179	cryptic gene	1.12	1.35		
	457646	AA725650	Hs.112948	ESTs	1.55	2.51		
	457733	AW974812	Hs.291971	ESTs	1.00	55.00		
50	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein	4.36	3.18		
	458092	BE545684	Hs.343566	KIAA0251 protein	1.00	1.32		
	458098	BE550224		metallothionein 1E (functional)	1.00	22.00		
	458207	T28472	Hs.7655	U2 small nuclear ribonucleoprotein auxi	2.06	1.88		
	458242	BE295588	Hs.28465	Homo sapiens cDNA: FLJ21869 fis, clone H	1.00	1.00		
55	458247	R14439	Hs.209194	ESTs	7.00	9.85		
	458679	AW975460	Hs.142913	ESTs	1.00	3.00		
	458778	AW451034	Hs.326525	arylsulfatase D	1.31	2.01		
	458933	AI638429	Hs.24763	RAN binding protein 1	1.98	1.71		
	459352	AW810383	Hs.206828	ESTs	12.60	63.00		
60	459670	FO1020	Hs.172004	titin	1.00	1.00		
	459702	AI204995		gb:an03c03.x1 Stratagene schizo brain S1	1.00	237.00		

TABLE 9B

65	Pkey:	Unique Eos probeset identifier number
	CAT number:	Gene cluster number
	Accession:	Genbank accession numbers
70	Pkey	CAT Number
	407746	10125_1
		AK001962 R69415 BE464605 AA418699 AA053293 AA149075 AA058396 AW338226 AW272859 AA454607 AI139535 AW469852 AI275461
		AW271982 AA730033 AA576507 AA991217 AA782067 AI985851 AA805864 AA505698 AW469857 R69546 AA988279 AW001647 N63320
		DB2661 T27343 AA306950 AA360989 R58778
		AW148852 BE350895
75	408070	1036688_1
	408660	107294_1
	409522	113735_1
	409866	1156522_1
	410032	1170436_1
	411089	123172_1
80	411152	1234028_1
	412537	1304_1
		AA456454 AA713730 AA091294 AA584921 N86077 AW836781 AA601031 AA579876 AA551106 AA633188 AW905577 AI955808 AI679386
		AI679895 AA514764 AA454562 AI082382 AA595822 AA551351 AA586369 AA666384 AA188934 AA666398 AA551297 AA565188
		BE069199 AW936012 AW877466 AW819782 AW935798 AW835546 AW936042 BE069121 AW835625 AW877536 AW935885 BE069202
		AW820019 AW935937 BE160180 AW935946 BE069101 BE069125 AW877527 BE160316 BE160398 AW935794 AW835701 AW935784
		AL031778 X59711 NM_002505 M59079 AI870439 AI494259 AW664010 AA405063 AA436132 BE174516 AA412691 AI400314 AA436024
		T29403 BE079412 BE079428 N90322 AI631202 AI141758 AI016793 AI167566 AI962075 AI375230 AI208445 AW235763 AL044113 AA382556
		AW953918 AA927051 AAB89823 BE003094 AW390155 AW360805 AW360823 AW360810 AA425472 AI694282 ALD44114 AI684577 AI809865

			A1478773 A1160445 A1674630 N69088 AW665529 A29278 A1129239 A1457890 A1621264 AW297152 A1268215 AA907787 A1286170 A1017982 A1963541 A1469807 A1969353 BE552356 N66509 AA736741 AA382555 AW075811 AW292026 H06382 AW957730 AA352014 R13591 AA121201 D60420 BE263253 BE047862 Z41952 A1424991 A1693507 A1863108 AA599060 A1091148 AA598689 R39887 AA813482 AW016452 H06383 R1807 A1364268 AA620528 A1241940 AW089149 AW090733 AW088875 Z38240 AA121202 R17734 BE157489 BE157560 AA826950 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245 AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R98127 T55938 BE279271 AW960304 T29812 AA476873 BE297387 AA292753 AA177048 NM_001826 X54941 BE314366 AA908783 A1719075 BE270172 BE269819 AA889955 A1204530 W25243 A1935150 AA872039 W72395 T99530 A1422691 H98460 N31428 BE255916 H03265 A1857576 AA776920 AA910644 AA459522 AA293140 AW514667 R75953 AW662396 AA662522 A1865147 A1423153 AW262230 AA584410 AA583187 AW024595 AW069734 A1828996 AA282997 AA876046 AW613002 AA527373 AW972459 A1831360 AA621337 AA100926 AA772418 AA594628 A1033892 W95096 A1034317 AA398727 A1085031 N95210 A1459432 A1041437 AA932124 AA627684 AA935829 A1004827 A1423513 A1094597 H42079 R54703 A1630359 AA617681 AA978045 AA643280 WA4561 A1991988 A1537692 A1090262 AA740817 A1312104 A1911822 AA416871 A1185409 AA129784 AA701623 A1075239 A1139549 AA633648 A1339996 A1336880 AA399239 A1078708 A1085351 A1362835 A146618 A1146955 A1989380 A1348243 N92892 AA765850 A1494230 A1278887 AA682596 A1492600 W80435 AA001979 R97424 A1129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785 A194211 AW059601 AW886710 R92790 N59755 A1361128 AW589407 H47725 H97534 H48076 H48450 T99631 AW300758 H03431 R76789 AA954344 H77576 R96823 AA57100 N92845 N49682 H42038 BE220698 BE220715 H95552 AA701624 N74173 R54704 H79520 H72923 H03265 BE261919 AA769633 AA480310 AA507454 AA910586 A1203723 AW104725 W25611 W25071 T89880 H03513 T77589 R9156 W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283872 A1267700 A1720344 AA191424 A1023543 A1469633 AA172056 AW958465 AA172236 AW953397 AA355086 AW265494 AA455904 AA195677 AW265432 AW991605 AA466370 N28754 N28747 A1568146 A1979339 AA322671 AA322672 AW955043 A1990326 AA776406 A1016250 AA843678 AW451882 N23137 N23129 W70051 A1038748 AA831327 A1925845 AW945895 Z42183 T31521 T97478 D62703 AA242966 D79798 A1076704 T74854 T74860 T72098 T73265 T73873 T69180 T74658 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74573 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M54982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61378 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R06796 H48353 T71914 T53939 T64121 AA633996 T72525 T67779 T68078 AA011465 AA343578 AV654847 AV654272 AV656001 A1084740 T82897 N33594 AA344542 AW805054 A1207457 T61743 AA026737 H84389 AA382695 AA918409 T68044 S82092 T39959 A1017721 AA312385 AA312919 T40156 H66239 AV652889 H38728 R85521 AV655200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343586 AW40774 AV651256 N54417 AA812862 AW182929 A111192 H61463 H72050 AA344503 H38639 A1277511 AV661108 A1207625 T7810 AA235252 T27853 T47778 R95746 H70620 AA701453 AW827166 R98475 C0825 AV657287 T71959 T71313 T73820 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72065 H72149 T73493 T73495 AV645993 T02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 A1110639 AA344603 AF063513 T64696 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74677 R23665 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA3428107 AA341964 T53747 T72042 T62764 A1064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 T29036 T72793 T69122 T64595 T62888 T69139 T68291 T64652 T67971 T46862 AA693592 A1248502 R29454 T64764 T57001 T73052 T71429 T61176 T58866 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068 A1792788 BE142230 AA252019 A1910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 A1571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 A1925552 AW950155 A191083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074058 AW009769 AW050690 AA858276 R55389 A1001051 AW050700 AW750216 AA614539 BE074045 A1307407 AW602303 BE073575 A1202532 AA524242 A1970839 A1909751 BE076078 A1909749 R55292 AW881145 AA490718 M56537 AA304575 T06067 AA331991 AL119930 AA320696 AW752565 A1031985 AL137241 A1792386 A1733664 A1857654 A1049911 AA337221 AA336756 AW966196 AW953120 R56325 AA349562 A193134 A1498691 AW771508 A1498457 A1768408 A1783624 A1383985 A1580267 D79813 AA393768 AK001536 AA191092 AW510354 A1554256 AL353968 AA134266 AA663848 AA400100 AA401424 A1038843 AA161338 BE268213 AA425597 N87306 AA092969 BE566038 AA247451 N47392 A1928802 AW182584 AW027872 A1819831 A1936994 W56258 A1653448 A1278611 A1283557 A1824306 AW338658 AW150899 AA687514 N47393 N29885 AA973469 A1038904 A1292064 A1034339 AW674593 N72156 A1079733 A1036883 A1291616 AA491599 AA93675 AA837380 BE006554 BE006473 A1087090 T33044 AA652043 A1203503 AA583959 W35283 A1129926 Z41844 AW020925 AW575848 A1684603 AA493297 A1140689 A1277176 AA425444 A1932767 W02632 BE396786 R37261 AW207206 AW341473 AA448195 A1951341 AA249027 A1038984 AK001993 A1080066 AV652725 BE566226 AA345557 AA315222 AA090585 AA375688 AA301092 AA298454 W05762 AW607939 A151658 D83880 N84323 BE296821 AW947007 D61461 AW079261 AA329482 AW901780 A1354442 AA772275 R31653 A1354441 A1767525 H92431 A1916735 H93575 A1394255 AW014741 A1573090 C06195 AW612857 AW265195 A1339558 A1377532 A1308821 A1919424 A1589705 AW055215 A1336532 A138051 AA806547 C75509 C00618 AW071172 AW769904 AA630381 A1678018 A1863985 D79662 BE221049 AW265018 A1589700 AW196655 N76573 A1370908 BE042393 N75017 A1698870 AW960115 AL133561 AL041090 AL117481 AL122069 AW439292 A1968826 AW072916 A1184913 AA489195 AW468994 AW469044 A59350 A1819642 A1280239 A1220572 AA789302 A1473611 AW841126 D60937 BE041395 AA491826 AA621946 AA715980 AA666102 AW970622 AA503009 AA502998 AA502989 AA502805 T92188 AA221036 R87170 BE537068 BE544757 C18935 AW812058 T92565 AA227415 AA233942 AA223237 AA668403 AA601627 AW869639 BE081833 BE000620 AW961170 AW847519 AA308542 AW821833 AW945688 C04699 AA205504 AA377241 AW821667 AA055720 AW817981 AW856468 AA155719 AA179928 T03007 AW754298 AA227407 AA113928 AA307904 C16859 A178376 S46400 AW811617 AW811616 W00557 BE142245 AW858232 AW861851 AW858362 AA232351 AA218567 AA055556 AW858231 AW857541 AW814172 H66214 AW814398 AF134164 AA243093 AA173345 AA199942 AA223384 AA227092 AA227080 T12379 AA092174 T61139 AA149776 AA698829 AW879188 AW813557 AW813538 A1267168 AA157718 AA157719 AA100472 AA100774 AA130756 AA157705 AA157730 AA157715 AA053524 AW849581 AW854566 C05254 AW882836 T92637 AW812621 AA208583 AA209204 BE156909 AA226824 A1829309 AW991957 N66951 AA527374 H66215 AA045564 A1694265 H60808 AA149726 AW195620 BE081333 BE073424 AW817662 AW817705 AW817703 AW817659 BE081531 H59570 AA628980 A126603 BE504035
--	--	--	--



438091 44964\_1 AW373052 T55662 AI299190 BE174210 AW579001 H01811 W40186 R67100 AI923886 AW952164 AA628440 AW896507 AW896516  
AA709126 AW898628 AW898644 AA947932 AW898625 AW898622 AI276125 AI185720 AW510698 AA987230 T52522 BE467708 AW243400  
AW043642 AI288245 AI186932 D52654 D55017 D52715 D52477 D53933 D54679 AI298739 AI146984 AI922204 N96343 BE174213 AA845571  
AI813854 AI214518 AI635262 AI139455 AI707807 AI698085 AW884528 AI024768 AI004723 AW087420 AI565133 N94964 AI268939  
AW513280 AI061126 AI435818 AI859106 AI360505 AI024767 AA513019 AA757598 X56198 AA902959 AI334784 AI860794 AA010207  
AW890091 AW513771 AI951391 AI337671 T52499 AA890205 AI640908 H75966 AA463487 AA358688 AI961767 AI866295 AA780994  
AI985913 BE174196 AA029094 AW592159 T55581 H79072 AI611201 AA910812 AI220713 AW149306 AI758412 AA045713 R79750 N76096  
AW979121 AA847986 AA829098  
AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882  
AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077  
AI109688 R23665 R26578  
AA570256 AW014761 AA573721 AI473237 AI022165 AA554071 AA127551 N90525 AW973623 AA447991 AA243852 BE328850 AI148171  
AI359627 AI005068 AI356567 AA232991 AW016855 AA906902 AA233101 AA127550 BE512923  
AI188710 AI032142 AW078833 N30308 AW675632 AI219028 AI341201 N22181 H95390  
W24187 W24194 R17789  
Y10043 NM\_005342 L05085 AL034450 BE614226 AW749053 AA379173 AA248230 BE514634 AA334622 R70656 AA367593 AA214649  
AA369318 AW957081 R05760 AA039903 AI886597 AW630122 AA906264 AA041527 R01145 AI088688 BE463637 AA398795 AI354883  
AI768938 AI569996 AI452952 AI168582 AI189869 AI066670 AW262560 AW613854 AA862839 AA435840 AA670197 AI024032 AI990859  
AI990069 N81095 AA847919 AW960150 AA211075 AA044704 AA367594 AW582587 AW858854 AW818630 AW818281 AW818433 AW582595  
AA096002 N83992  
AI71630 BE540637 BE265481 AW407710 BE513882 BE546739 AA053597 BE140503 BE218514 AW956702 AI656234 AI636283 AI567265  
AW340858 BE207794 AA053085 R69173 AA292343 AA454908 AA293504 AI659741 AI927478 AA399460 AI760441 AA346416 BE047245  
AA730380 AA394063 AA454833 AI982791 AI567270 AI813332 AI767858 AA427705 D20284 AI221458 BE048537 AI263048 AA346417  
AA911497 BE537702  
AI638293 AW813561  
AI761324 AW880941 AW880937  
AW118072 AI631982 T15734 AA224195 AI701458 W20198 F26326 AA890570 N90552 AW071907 AI671352 AI375892 T03517 R88265  
AI124088 AA224388 AI084316 AI354686 T33652 AI140719 AI720211 T03490 AI372637 T15415 AW205836 AA630384 T03515 T33230  
AA017131 AA443303 T33623 AI222556 T33511 T33785 AI419606 D55612  
W52854 AL117600 BE208116 BE208432 BE206239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291  
AW450652 AW449519 AA933634 AI808539 AA351618 AW449522 AI827626 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756  
AI133619 AA468118 AA383064 AI476447 T09430 AI673758 AA524895 AI581345 AI300820 AW498812 AA256162 AI559724 AI685732  
AA602400 AA905453 AI204595 AW166541 AA157456 AA156269 AA383652 AA431072 AW592707 AI435410 AW272464 AI215594 AA622747  
R74039 N35031 AI804128 AW513621 AA868351 AI026826 AI493388 AA614641 W81604 AI557080 AI214351 AA730140 AI125754 AI200813  
AI269603 AI565082 AI807095 AI476629 AA505909 AI368449 AI686077 AI582930 AW085038 AA757863 AA730154 AI767072 AA466816  
AI734130 AI734138 AA426284 AA433997 AI741241 AW043563 AI732741 AI732734 AA437369 AA425820 AA664048 R74130  
BE144666 BE184942 AW238414 BE184946  
AW993247 AW861464  
AA203682 R11958  
BE550224 AA832519 N45402 AW885857 N29245 BE455409 W07677 AW970089 AI299731 AA482971 BE503548 H18151 W79223 AF086393  
AA461301 W74510 R34182 AI090689 N46003 BE071550 R28075 AW134982 AI240204 AI138906 AW026179 AI572316 BE466182 AI206395  
AI276154 AI273269 AI422817 AI371014 AI421274 AI188525 AA939164 BE549810 AW137865 AI694998 BE503841 AA459718 BE327407  
BE467534 BE218421 BE467757 AA989054 BE467063 AI797130 BE327781

TABLE 9C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1995) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
Nt\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400512	9796593	Minus	1439-1615
400517	9796688	Minus	49996-50346
400560	9843598	Plus	94182-94323,97056-97243,101095-101236,102824-103005
400664	8118496	Plus	13558-13721,13942-14090,14554-14679
400665	8118496	Plus	16879-17023
400666	8118496	Plus	17982-18115,20297-20456
400749	7331445	Minus	9162-9293
400763	8131616	Minus	35537-35784
401027	7230983	Minus	70407-70554,71060-71160
401093	8516137	Minus	22335-23166
401203	9743387	Minus	172961-173056,173868-173928
401212	9858408	Plus	67839-88028
401411	7799787	Minus	144144-144329
401435	8217934	Minus	54508-55233
401464	6682291	Minus	170688-170834
401714	6715702	Plus	96484-96681
401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131258,131866-131932,132451-132575,133580-134011 83126-83250,85320-85540,94719-95287 28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573 83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,86290-86814 165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942 6973-7118 124054-124209 61542-61750 42904-43124,43211-43336,44607-44763,45199-45281,46337-46732 121907-122035,122804-122921,124019-124161,124455-124610,125672-126076 113765-113910,115653-115765,116808-116940 21059-21168 35279-35405,35573-35659 110326-110491
401760	9929699	Plus	
401780	7249190	Minus	
401781	7249190	Minus	
401785	7249190	Minus	
401797	6730720	Plus	
401961	4581193	Minus	
401985	2580474	Plus	
401994	4153858	Minus	
402075	8117407	Plus	
402260	3399665	Minus	
402265	3287673	Plus	
402297	6598824	Plus	
402408	9796239	Minus	



	402420	9795339	Plus	129750-129919
	402674	8077108	Minus	39280-39502
	402802	3287156	Minus	53242-53432
5	402994	2956643	Minus	4727-4969
	403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
	403306	8099945	Plus	127100-127251
	403329	8516120	Plus	96450-96598
	403381	9438267	Minus	26009-26178
10	403478	9958258	Plus	118458-118564
	403485	9966528	Plus	2888-3001,3198-3532,3855-4117
	403627	8569879	Minus	23868-24342
	403715	7239669	Plus	85128-85292
	404044	9558573	Minus	225757-225939
15	404076	9931752	Minus	3848-3967
	404101	8076925	Minus	125742-125997
	404140	9843520	Plus	37761-38147
	404165	9926489	Minus	69025-69128
	404185	4572584	Minus	129171-129327
20	404210	6006246	Plus	169926-170121
	404253	9367202	Minus	55675-55055
	404287	2326514	Plus	53134-53281
	404298	9944263	Minus	73591-73723
	404347	9838195	Plus	74493-74829
25	404440	7528051	Plus	80430-81581
	404721	9856548	Minus	173763-174294
	404794	4826439	Plus	101619-101898
	404854	7143420	Plus	14260-14537
	404877	1519284	Plus	1095-2107
30	404927	7342002	Plus	68690-69663
	404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40874,42351-42450
	405449	7622497	Plus	42236-42570
	405568	6006906	Plus	35912-36065
	405572	3800891	Plus	85230-85938
35	405646	4914350	Plus	741-969
	405676	4557087	Plus	73195-73917
	405770	2735037	Plus	61057-62075
	405832	7787812	Minus	123525-123713
40	406137	9165422	Minus	30487-31058
	406360	9256107	Minus	7513-7673
	406399	9256288	Minus	63448-63554
	406467	9795551	Plus	182212-182958

TABLE 10A: Potential Therapeutic, Diagnostic and Prognostic targets for Therapy of Lung Cancer and Non-malignant Lung Disease

Table 2A shows about 307 genes up-regulated in non-malignant lung disease relative to lung tumors and normal body tissues and/or down-regulated in lung tumors relative to normal lung and non-malignant lung disease. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 10B show the accession numbers for those Pkey's lacking UnigeneID's for table 10A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 10C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 10A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

	Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
65	404394			ENSP00000241075:TRRAP PROTEIN.	0.79	3.10
	404916			Target Exon	1.00	159.00
	405257			Target Exon	1.00	422.00
70	407228	M25079	Hs.155378	hemoglobin, beta	0.47	2.33
	407568	AA740964	Hs.62699	ESTs	1.00	123.00
	408562	AI438323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
	409031	AA376836	Hs.76728	ESTs	1.00	128.00
	410434	AF051152	Hs.63668	tol-like receptor 2	39.65	149.00
75	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	1.00	109.00
	410808	T40326	Hs.167793	ESTs	1.14	13.14
	412351	AL135860	Hs.73828	T-cell acute lymphocytic leukemia 1	0.37	2.27
	412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
	413795	AL040178	Hs.142003	ESTs	0.10	11.90
80	414154	AW205314	Hs.323060	ESTs	0.62	2.09
	414214	D49958	Hs.75819	glycoprotein M5A	0.03	4.55
	414998	NM_002543	Hs.77729	oxidised low density lipoprotein (lectin	0.64	2.97
	415122	D60708	Hs.22245	ESTs	0.07	8.97
	415765	NM_005424	Hs.78824	tyrosine kinase with immunoglobulin and	0.67	1.65
85	415775	H00747	Hs.29792	ESTs, Weakly similar to 138022 hypothe	0.29	2.64
	415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00

	416319	AJ815601	Hs.79197	CD83 antigen (activated B lymphocytes, I	15.32	237.00
	416402	NM_000715	Hs.1012	complement component 4-binding protein,	0.64	4.00
	417355	D13188	Hs.82002	endothelin receptor type B	0.01	3.90
5	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	38.30	357.00
	417511	AL049176	Hs.82223	chordin-like	1.00	179.00
	418489	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	0.02	6.00
	418726	BE241812	Hs.87850	protein tyrosine phosphatase, non-recept	1.00	113.00
	418741	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom	0.44	1.90
10	418883	BE387036	Hs.1211	acid phosphatase 5, tartrate resistant	0.96	2.04
	419086	NM_000216	Hs.89591	Kallmann syndrome 1 sequence	0.62	2.74
	419150	T29618	Hs.89640	TEK tyrosine kinase, endothelial (venous	0.03	6.90
	419235	AW470411	Hs.288433	neurotrophin	1.48	5.13
	419407	AW410377	Hs.41502	hypothetical protein FLJ21276	37.55	336.00
	420556	AA278300	Hs.124292	Homo sapiens cDNA: FLJ23123 fs, clone L	0.80	3.65
15	420656	AA279098	Hs.187636	ESTs	1.65	8.07
	420729	AW964897	Hs.290825	ESTs	2.99	25.82
	421177	AW070211	Hs.102415	Homo sapiens mRNA; cDNA DKFZp586N0121 (f	0.46	1.95
	422060	R20893	Hs.325823	ESTs, Moderately similar to ALU5_HUMAN A	1.00	156.00
	422426	W79117	Hs.58559	ESTs	0.03	7.44
20	422652	AW957969	Hs.118958	syntaxin 11	0.14	3.62
	423099	NM_002837	Hs.123641	protein tyrosine phosphatase, receptor t	0.01	3.16
	424433	H04607	Hs.9218	ESTs	0.75	141.75
	424585	AA464840	Hs.131987	ESTs	1.00	167.00
25	424711	NM_005795	Hs.152175	calcitonin receptor-like	0.43	3.01
	424973	X92521	Hs.154057	matrix metalloproteinase 19	0.37	19.45
	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipi	0.14	3.35
	425664	AJ066276	Hs.159003	transient receptor potential channel 6	1.00	94.00
	425998	AJ076629	Hs.165950	fibroblast growth factor receptor 4	0.68	1.42
30	426657	NM_015865	Hs.171731	solute carrier family 14 (urea transport	0.03	3.74
	426753	T89832	Hs.170278	ESTs	1.00	141.00
	427558	D49493	Hs.2171	growth differentiation factor 10	1.00	117.00
	427983	M17706	Hs.22233	colony stimulating factor 3 (granulocyte	0.75	2.20
	428467	AK002121	Hs.184465	hypothetical protein FLJ11259	0.76	2.25
35	428927	AA441837	Hs.90250	ESTs	0.01	3.62
	429496	AA453800	Hs.192793	ESTs	1.00	138.00
	430468	NM_004673	Hs.241519	angiotensin-like 1	1.00	132.00
	431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
	431728	NM_007351	Hs.268107	multimerin	1.00	157.00
40	431848	AJ378857	Hs.128758	ESTs, Highly similar to AF175283 1 zinc	0.34	2.24
	432128	AA127221	Hs.117037	ESTs	0.00	1.15
	432519	AJ221311	Hs.130704	ESTs, Weakly similar to BCHUIA S-100 pro	0.01	2.06
	433043	W57554	Hs.125019	lymphoid nuclear protein (LAF-4) mRNA	1.00	267.00
	433803	AJ823593	Hs.27688	ESTs	1.00	105.00
45	434730	AA644669	Hs.193042	ESTs	1.05	3.15
	435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	0.83	1.94
	436532	AA721522		gb:nv54h12r1 NCL_CGAP_Ew1 Homo sapiens	1.00	218.00
	437119	AJ379921	Hs.177043	ESTs	1.00	133.00
	437140	AA312789	Hs.283689	activator of CREM in testis	0.67	122.67
50	437211	AA382207	Hs.5509	ecotropic viral integration site 2B	1.00	142.00
	437960	AJ669586	Hs.222194	ESTs	1.00	147.00
	438202	AW169287	Hs.22588	ESTs	1.00	141.00
	438873	AJ302471	Hs.124292	Homo sapiens cDNA: FLJ23123 fs, clone L	0.71	3.66
	438875	AA827640	Hs.189059	ESTs	23.32	370.00
55	441048	AA913488	Hs.192102	ESTs	0.77	8.50
	441188	AW292830	Hs.255609	ESTs	3.43	16.36
	441499	AW296235	Hs.101689	ESTs	1.00	167.00
	444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
	444527	NM_005408	Hs.11383	small inducible cytokine subfamily A (Cy	46.47	153.00
60	444561	NM_004469	Hs.11392	c-fos induced growth factor (vascular en	0.01	3.08
	445279	R41900	Hs.22245	ESTs	0.60	141.00
	446017	N98238	Hs.55185	ESTs	0.18	2.39
	446984	AB020722	Hs.16714	Rho guanine exchange factor (GEF) 15	0.10	2.16
	446998	N99013	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 (f	0.01	2.53
65	447357	AJ375922	Hs.159367	ESTs	0.46	2.64
	448106	AJ800470	Hs.171941	ESTs	18.05	296.00
	448253	H25899	Hs.201591	ESTs	1.00	141.00
	449275	AW450848	Hs.205457	perlecan	0.56	1.38
	450400	AJ694722	Hs.279744	ESTs	0.88	4.33
70	450696	AJ654223	Hs.16026	hypothetical protein FLJ23191	0.52	2.08
	450726	AW204600	Hs.250505	retinoic acid receptor, alpha	0.79	2.01
	451497	H83294	Hs.284122	Wnt inhibitory factor-1	0.35	2.03
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidy)	0.13	2.25
	453636	R67837	Hs.169872	ESTs	1.00	116.00
75	458332	AJ000341	Hs.220491	ESTs	1.00	192.00
	459580	AA022888	Hs.176065	ESTs	0.20	2.98
	400269			Eos Control	0.40	2.40
	403421			NM_016369: Homo sapiens claudin 18 (CLDN	0.53	1.77
	407570	Z19002	Hs.37096	zinc finger protein 145 (Kruppel-like, e	0.01	3.18
80	412295	AW068826	Hs.117176	poly(A)-binding protein, nuclear 1	0.56	1.74
	414517	M24461	Hs.76305	surfactant, pulmonary-associated protein	0.64	1.50
	417204	N81037	Hs.1074	surfactant, pulmonary-associated protein	0.33	1.16
	418307	U70867	Hs.83974	solute carrier family 21 (prostaglandin	0.53	1.55
	418935	T28499	Hs.89485	carbonic anhydrase IV	0.20	1.28
85	421502	AF111856	Hs.105039	solute carrier family 34 (sodium phospho	0.78	1.90
	421798	N74880	Hs.29877	N-acylsphingosine amidohydrolase (acid c	0.59	1.54

	423354	AB011130	Hs.127436	calcium channel, voltage-dependent, alpha	0.59	1.55
	423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	0.35	1.62
5	425438	T62216	Hs.270840	ESTs	0.23	9.45
	426828	NM_000020	Hs.172570	activin A receptor type II-like 1	0.03	1.71
	427019	AA001732	Hs.173233	hypothetical protein FLJ10970	0.01	1.49
	428043	T92248	Hs.2240	uteroglobin	0.42	1.26
	430280	AA381258	Hs.237868	interleukin 7 receptor	0.46	2.43
10	431433	X65018	Hs.253495	surfactant, pulmonary-associated protein	0.57	1.59
	431723	AW058350	Hs.16762	Homo sapiens mRNA; cDNA DKFZp564B2062 f	0.29	1.80
	432985	T92363	Hs.178703	ESTs	0.32	2.27
	441835	AB036432	Hs.184	advanced glycosylation end product-speci	0.31	1.51
	442275	AW449467	Hs.54795	ESTs	0.55	1.78
15	443709	AI082692	Hs.134662	ESTs	0.00	3.02
	444325	AW152618	Hs.16757	ESTs	0.32	2.49
	450954	AI904740	Hs.25691	receptor (calcionin) activity modifying	0.46	1.74
	451558	NM_001089	Hs.26630	ATP-binding cassette, sub-family A (ABC1	0.52	1.87
	453310	X70697	Hs.553	solute carrier family 8 (neurotransmitte	0.00	3.30
	456855	AF035528	Hs.153863	MAD (mothers against decapentaplegic, Dr	0.01	2.31
20	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	0.66	2.20
	400754			Target Exon	1.00	297.00
	401045			C11001883*:g 6753278 ref NP_033938.1  c	1.00	109.00
	401083			NM_016582*:Homo sapiens peptide transpor	0.89	1.39
25	402474			NM_004079:Homo sapiens cathepsin S (CTSS	1.45	4.47
	402808			ENSP00000235229:SEMBL	1.00	1.87
	403021			C21000030:g 9955960 ref NP_063957.1  AT	1.00	149.00
	403438			NM_031419*:Homo sapiens molecule possess	1.06	2.96
	403687			NM_007037*:Homo sapiens a disintegrin-li	0.04	4.89
30	403764			NM_005463:Homo sapiens heterogeneous nuc	1.00	225.00
	404277			NM_019111*:Homo sapiens major histocompa	0.97	1.93
	404288			NM_002944*:Homo sapiens v-ros avian UR2	1.00	68.00
	404518	AI815601		CD83 antigen (activated B lymphocytes, l	0.02	1.83
	405106			C11001637*:g 5032241 ref NP_005732.1  z	1.00	235.00
35	405381			Target Exon	1.00	93.00
	406387			Target Exon	1.37	6.02
	406646	M33600		major histocompatibility complex, class	0.86	2.46
	406714	AI219304	Hs.266959	hemoglobin, gamma G	0.01	3.19
40	406753	AA505665	Hs.217493	annexin A2	1.00	147.00
	406973	M34996	Hs.198253	major histocompatibility complex, class	1.03	2.04
	407248	U82275	Hs.94498	leukocyte immunoglobulin-like receptor,	1.00	64.00
	407510	U96191		gb:Human trophoblast hypoxia-regulated f	1.00	90.00
	407731	NM_000066	Hs.38069	complement component 8, beta polypeptide	1.00	67.00
	407830	NM_001086	Hs.587	arylacetamide deacetylase (esterase)	1.00	102.00
45	408045	AW138959	Hs.245123	ESTs	1.00	70.00
	408074	R20723		ESTs	1.00	112.00
	408374	AW025430	Hs.155591	forkhead box F1	0.07	10.17
	409064	AA062954	Hs.141883	ESTs	0.39	2.31
	409083	AF050083	Hs.673	interleukin 12A (natural killer cell sti	1.00	95.00
50	409153	W03754	Hs.50813	hypothetical protein FLJ20022	0.01	4.55
	409203	AA780473	Hs.687	cytochrome P450, subfamily IVB, polypept	0.01	3.72
	409238	AL049990	Hs.51515	Homo sapiens mRNA; cDNA DKFZp564G112 (fr	1.00	79.00
	409389	AB007979	Hs.301281	Homo sapiens mRNA, chromosome 1 specific	0.14	27.35
	409718	D86840	Hs.56045	src homology three (SH3) and cysteine ri	1.00	113.00
55	409798	BE178622	Hs.16291	gb:PM3-HT0605-270200-001-a02 HT0605 Homo	0.64	2.47
	411020	NM_006770	Hs.67726	macrophage receptor with collagenous str	0.55	2.40
	411687	BE160198		gb:QV1-HT0413-010200-059-h03 HT0413 Homo	1.00	111.00
	412000	AW576555	Hs.15780	ATP-binding cassette, sub-family A (ABC1	1.00	95.00
	412358	BE047490	Hs.24172	ESTs	1.00	87.00
60	412420	AL035668	Hs.73853	bone morphogenetic protein 2	1.43	8.07
	412564	X83703	Hs.31432	cardiac ankyrin repeat protein	0.02	3.07
	412869	AA290712	Hs.82407	CXC chemokine ligand 16	0.93	1.72
	412870	N22788	Hs.82407	CXC chemokine ligand 16	0.97	1.51
	413529	U11874	Hs.846	interleukin 8 receptor, beta	0.02	2.42
65	413533	BE146973		gb:QV4-HT0222-011199-019-e05 HT0222 Homo	0.65	1.50
	413689	BE157286	Hs.20631	zinc finger protein, subfamily 1A, 5 (Pe	20.87	232.00
	413724	AA131466	Hs.23767	hypothetical protein FLJ12666	1.00	80.00
	413800	AI129238	Hs.192235	ESTs	1.00	85.00
	413802	AW964490	Hs.32241	ESTs, Weakly similar to S65657 alpha-1C-	1.00	213.00
70	413829	NM_001872	Hs.75572	carboxypeptidase B2 (plasma)	0.02	3.93
	414376	BE393856	Hs.66915	ESTs, Weakly similar to 16.7Kd protein (	1.00	115.00
	414577	AI056548	Hs.72116	hypothetical protein FLJ20992 similar to	0.49	1.94
	414700	H63202	Hs.38163	ESTs	0.03	3.75
	415078	AA311223	Hs.283091	found in inflammatory zone 3	0.66	1.95
75	415120	N84464	Hs.34950	ESTs	1.00	120.00
	415323	BE269352	Hs.949	neutrophil cytosolic factor 2 (65kD, chr	0.60	2.48
	415335	AA847758	Hs.111030	ESTs	1.00	95.00
	415582	W92445	Hs.165195	Homo sapiens cDNA FLJ14237 fis, clone NT	1.00	136.00
	416030	H15261	Hs.21948	ESTs	0.02	8.07
80	416427	BE244050	Hs.79307	Rac/Cdc42 guanine exchange factor (GEF)	1.00	73.00
	416464	NM_000132	Hs.79345	coagulation factor VIII, procoagulant co	0.70	3.36
	416585	X54162	Hs.79386	leiomodin 1 (smooth muscle)	0.06	6.56
	416847	L43821	Hs.80261	enhancer of filamentation 1 (cas-like do	0.70	3.66
	417148	AA359896	Hs.293885	hypothetical protein FLJ14902	1.00	114.00
85	417370	T28551	Hs.82030	tryptophanyl-tRNA synthetase	0.85	1.30
	417673	T87281	Hs.16355	ESTs	0.15	15.54

5	418067	Al127958	Hs.83393	cystatin E/M	0.81	1.74
	418296	C01566	Hs.86671	ESTs	1.00	99.00
	418643	J03798	Hs.86948	small nuclear ribonucleoprotein D1 polyp	1.00	60.00
	418832	X04011	Hs.88974	cytochrome b-245, beta polypeptide (chro	2.40	14.74
	418945	BE246762	Hs.89499	arachidonate 5-lipoxygenase	0.67	3.16
	419261	X07876	Hs.89791	wingless-type MMTV integration site fami	1.00	73.00
	419564	U08989	Hs.91139	solute carrier family 1 (neuronal/epitha	1.00	192.00
	419574	AK001989	Hs.91165	hypothetical protein	1.00	94.00
10	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	61.16	500.00
	420256	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula	0.52	1.70
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	1.00	172.00
	420577	AA278436	Hs.186649	ESTs	1.00	97.00
	421262	AA286746	Hs.9343	Homo sapiens cDNA FLJ14265 fis, clone PL	1.00	64.00
15	421445	AA913059	Hs.104433	Homo sapiens, clone IMAGE:4054868, mRNA	0.88	1.51
	421470	R27496	Hs.1378	annexin A3	0.05	11.26
	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	1.00	73.00
	421563	NM_006433	Hs.105806	granulysin	0.82	2.42
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	5.50	31.57
20	421855	F06504	Hs.27384	ESTs, Moderately similar to ALU4_HUMAN A	1.00	129.00
	421913	AI934365	Hs.109439	osteoglycin (osteoinductive factor, mime	1.00	101.00
	421952	AA300900	Hs.98849	ESTs, Moderately similar to AF161511 H	0.60	63.60
	422232	D43945	Hs.113274	transcription factor EC	1.00	148.00
	422386	AF105374	Hs.115830	heparan sulfate (glucosamine) 3-O-sulfot	1.40	3.98
25	423168	R34385	Hs.124940	GTP-binding protein	0.34	3.59
	423196	AK001866	Hs.125139	hypothetical protein FLJ11004	0.55	2.00
	423387	AJ012074		vasoactive intestinal peptide receptor 1	0.09	2.13
	423424	AF150241	Hs.128433	prostaglandin D2 synthase, hematopoietic	1.00	141.00
	423456	AL110151	Hs.128797	DKFZP586D0824 protein	1.00	66.00
30	423696	Z92546		Sushi domain (SCR repeat) containing	0.73	1.27
	424027	AW337575	Hs.201591	ESTs	0.54	2.58
	424212	NM_005814	Hs.143131	glycoprotein A33 (transmembrane)	0.77	2.47
	425087	R62424	Hs.126059	ESTs	1.00	74.00
	425175	AF020202	Hs.155001	UNC13 (C. elegans)-like	0.85	1.96
35	425771	BE561776	Hs.159494	Bruton agammaglobulinemia tyrosine kinas	1.18	2.56
	426486	BE172885	Hs.170056	Homo sapiens mRNA; cDNA DKFZp586B0220 (f	1.00	76.00
	427507	AF240467	Hs.179152	tol-like receptor 7	1.00	63.00
	427618	NM_000760	Hs.2175	colony stimulating factor 3 receptor (gr	0.60	2.19
	427732	NM_002980	Hs.2199	secretin receptor	0.97	1.42
40	427952	AA765368	Hs.293941	ESTs, Moderately similar to A63959 throm	1.00	105.00
	428709	BE268717	Hs.104916	hypothetical protein FLJ21940	1.00	80.00
	428769	AW207175	Hs.106771	ESTs	0.09	2.55
	428780	AI478578	Hs.50636	ESTs	1.00	98.00
	428833	AI928355	Hs.185805	ESTs	1.00	113.00
45	429657	D13626	Hs.2465	KIAA0001 gene product putative G-protei	1.00	52.00
	430212	AA469153		gbnc67f04.s1 NCL CGAP_Prl Homo sapiens	1.00	132.00
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	0.11	15.60
	430376	AW292053	Hs.12532	chromosome 1 open reading frame 21	1.00	103.00
50	430414	AW365565	Hs.120388	ESTs	0.50	6.96
	430656	AA482900	Hs.162080	ESTs	1.00	70.00
	430843	AI734149	Hs.119514	ESTs	1.00	90.00
	430998	AF128847	Hs.204038	indolethylamine N-methyltransferase	0.29	1.84
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	1.00	79.00
	431921	N46466	Hs.58879	ESTs	0.91	1.67
55	432176	AW090386	Hs.112278	arrestin, beta 1	0.66	2.63
	432203	AA305746	Hs.49	macrophage scavenger receptor 1	1.00	76.00
	432231	AA339977	Hs.274127	CLST 11240 protein	0.46	1.46
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	0.79	2.25
	432522	D11466	Hs.51	phosphatidylinositol glycan, class A (pa	1.93	4.83
60	432596	AJ224741	Hs.278461	mainrin 3	0.04	5.79
	432850	X87723	Hs.3110	angiotensin receptor 2	1.00	167.00
	433138	AB029496	Hs.59729	semaphorin sem2	0.04	9.16
	433563	AI732637	Hs.277901	ESTs	1.00	91.00
	433588	AI056872	Hs.133386	ESTs	120.16	315.00
65	434445	AI349306	Hs.11782	ESTs	0.60	1.84
	435496	AW840171	Hs.265398	ESTs, Weakly similar to transformation-r	1.00	128.00
	435974	U29690	Hs.37744	Homo sapiens beta-1 adrenergic receptor	1.00	108.00
	436061	AI248584	Hs.190745	Homo sapiens cDNA: FLJ21326 fis, clone C	1.00	91.00
	437157	BE048860	Hs.120655	ESTs	1.00	87.00
70	437207	T27503	Hs.15929	hypothetical protein FLJ12910	1.00	105.00
	437311	AA370041	Hs.9456	SWI/SNF related, matrix associated, acti	1.00	71.00
	437439	H29796	Hs.269622	ESTs	1.00	115.00
	438199	AW016531	Hs.122147	ESTs	1.00	80.00
	439551	W72062	Hs.11112	ESTs	0.30	3.10
75	440515	AJ131245	Hs.7239	SEC24 (S. cerevisiae) related gene fami	1.00	77.00
	440887	AI799488	Hs.135905	ESTs	1.00	85.00
	441025	AA913880	Hs.176379	ESTs	1.00	82.00
	441384	AA447849	Hs.288660	Homo sapiens cDNA: FLJ22182 fis, clone H	0.79	1.89
	441735	AI738675	Hs.127346	ESTs	1.00	75.00
80	442200	AW590572	Hs.235768	ESTs	0.78	5.83
	442832	AW206560	Hs.253569	ESTs	0.03	10.88
	442957	AI849952	Hs.49397	ESTs	1.00	70.00
	443282	T47764	Hs.132917	ESTs	1.00	197.00
	443547	AW271273	Hs.23767	hypothetical protein FLJ12666	1.00	253.00
85	443851	F13272	Hs.111334	ferritin, light polypeptide	0.55	2.09
	444330	AI597655	Hs.49265	ESTs	1.00	90.00

	444515	AW204508	Hs.169979	ESTs	1.00	84.00
	445769	AI714171	Hs.23666	ESTs	0.02	4.38
	445908	R13580	Hs.13436	Homo sapiens clone 24425 mRNA sequence	1.00	97.00
	446291	BE397753	Hs.14623	interferon, gamma-inducible protein 30	0.93	1.69
5	446917	AI347863	Hs.156672	ESTs	1.00	106.00
	447261	NM_006691	Hs.17917	extracellular link domain-containing 1	0.40	47.20
	447432	AW958473	Hs.301957	nudix (nucleoside diphosphate linked mol	1.00	100.00
	447482	AB033059	Hs.18705	KIAA1233 protein	0.05	8.21
10	447997	H00656	Hs.29792	ESTs, Weakly similar to I38022 hypothetical	0.02	5.42
	448299	AA497044	Hs.20887	hypothetical protein FLJ10392	1.00	79.00
	448782	AL050295	Hs.22039	KIAA0758 protein	0.42	1.56
	450575	NM_005859	Hs.29117	purine-rich element binding protein A	0.17	11.33
	450584	AA040403	Hs.60371	ESTs	1.00	94.00
15	450693	AW450461	Hs.203965	ESTs	1.00	91.00
	450715	AI266484	Hs.31570	ESTs, Weakly similar to KIAA1324 protein	1.00	152.00
	451103	R52804	Hs.25956	DKFZP564D206 protein	1.00	68.00
	451220	AF124251	Hs.26054	novel SH2-containing protein 3	0.60	1.30
	451668	Z43948	Hs.326444	cartilage acidic protein 1	0.54	1.91
	452197	AW023595	Hs.232048	ESTs	1.00	67.00
20	452331	AA598509	Hs.28117	purine-rich element binding protein A	4.53	11.07
	452353	C18825	Hs.29191	epithelial membrane protein 2	0.72	2.24
	453049	BE537217	Hs.30343	ESTs	1.00	68.00
	453107	NM_016113	Hs.279746	vanilloid receptor-like protein 1	0.83	1.70
25	453355	AW295374	Hs.31412	Homo sapiens cDNA FLJ11422 fis, clone HE	1.00	132.00
	453390	AA862496	Hs.28482	ESTs	1.00	72.00
	453631	AA417940		ESTs, Weakly similar to JC5795 CDEP prot	1.00	68.00
	454741	BE154396		gb:CM2-HT0342-091299-050-b05 HT0342 Homo	0.57	2.89
	456579	AA287827	Hs.284205	up-regulated by BCG-CWS	1.00	82.00
30	456672	AK002016	Hs.114727	Homo sapiens, clone MGC:16327, mRNA, com	0.79	1.96
	457400	AF032906	Hs.252549	cathepsin Z	1.03	3.25
	457718	F18572	Hs.22978	ESTs, Weakly similar to ALU4_HUMAN ALU S	1.00	113.00
	459696	F03027		gb:HSC1KA072 normalized Infant brain cDN	1.00	544.00

TABLE 10B

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
408074	103684_1	R20723 AA263003 AA333976 AA334725 AA334151 AW965490 AA310513 AJ810530 D31302 AW134897 AA830127 AA046953 AI668930 CO6094 AW104534
411667	1253334_1	BE160198 AW935898 T11520 AW835930 AW856073 AW861034
413533	1375344_1	BE146973 BE146972 BE147042 BE147018 BE145783 BE147020 BE146781 BE147019 BE146766 BE147021 BE146952 BE146767 BE147044 BE146797 BE146776 BE146985 BE146793 BE146768 BE146771 BE146954 BE146760 BE147048 BE147025 BE147030
423387	22779_1	AJ012074 U11087 L13288 X75299 L20295 AW630780 H14880 T28037 AJ872991 R72136 AW449839 T81622 T79697 T29519 R94105 T83923 R73300 AJ797007 R73390 AA961010 H74168 AI689932 BE045543 AI808418 AI608912 AI806573 AW884084 AW872978 AW872985 AA565655 AI022915 R50647 R73210 H45098 R46451 AW166269 T71132 AI264547 R52146 AI304920 R73391 AW884059 AW884085 H73241 T60038 T79612 R73145 R50549 AJ094557 AI668793 R72302 AJ564366 W01956 AA418962 W32571 R72840 H45409 R72085 R46356 R46758 AA508805 AA418798 T83751 R94072 T16182 AA928785 AA903896 Z92546 AA330586 AI570568 AW341487 AI827050 AW298668 AI792189 AI015693 AJ733599 AJ572251 AI672488 AW193262 AI244716 AI864375 AI206100 AA912444 AI269365 AI640254 AW772466 AI867336 AA627604 H16914 AA358477 AA338009
430212	314437_1	AA489153 AJ718503 AA469225
435532	421802_1	AA721522 AW975443 T93070
453531	97026_1	AA417940 AA036735 T07025
454741	1232559_1	BE154396 AW817959 BE154393

TABLE 10C

Pkey: Unique number corresponding to an Eos probeset  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA  
 sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 NT\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NT_position
400754	7331445	Plus	144559-144684
401045	8117619	Plus	90044-90184,91111-91345
401083	3242744	Plus	33192-33380
402474	7547175	Minus	53526-53828,55755-55920,57530-57757
402808	6456148	Minus	114964-115136,115461-115585,115931-116047,117666-117771,118004-118102
403021	7547270	Plus	120789-120966
403421	9565041	Minus	126609-126773,139986-140205
403438	9719579	Plus	90792-90938
403687	7387384	Plus	9009-9534
403764	7717105	Minus	118692-118853
404277	1834458	Minus	91665-91946
404288	2769844	Plus	3512-3691
404394	3135305	Minus	37121-37205,37491-37762,41053-41140,41322-41593,41773-41919
404518	8151988	Plus	84494-84603
404916	7341826	Plus	91057-91188
405106	8079395	Minus	80877-81418
405257	7329310	Plus	73121-73273
405381	6006920	Minus	7636-8054

**WO 02/086443**  
406387 9256180 Plus

116229-116371,117512-117651

**PCT/US02/12476**

TABLE 11A: Genes Distinguishing Adenocarcinoma from Other Lung Diseases and Normal Lung

Table 11A shows about 84 genes upregulated in lung adenocarcinomas relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymatrix Hu03 Genechip array.

Table 11B show the accession numbers for those Pkey's lacking UnigeneID's for table 11A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 11C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 11A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset Identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
403329			Target Exon	1.00	61.00
406399			NM_003122*Homo sapiens serine protease	1.00	39.00
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	226.37	350.00
407869	A1827976	Hs.24391	hypothetical protein FLJ13612	0.77	1.18
407881	AW072003	Hs.40958	heparan sulfate (glucosamine) 3-O-sulfot	1.00	10.00
408908	BE296227	Hs.250822	serine/threonine kinase 15	7.76	1.00
409103	AF251237	Hs.112208	XAGE-1 protein	80.44	40.00
409187	AF154830	Hs.50956	carbamoyl-phosphate synthetase 1, mitoch	1.00	1.00
409269	AA576953	Hs.22972	hypothetical protein FLJ13352	1.00	1.00
410076	T05387	Hs.7991	ESTs	1.12	1.50
410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fis, clone HE	9.89	1.00
410399	BE068889		synuclein, gamma (breast cancer-specific	0.92	1.06
411908	L27943	Hs.72924	cytidine deaminase	1.00	1.00
412612	NM_000047	Hs.74131	arylsulfatase E (chondrodysplasia puncta	1.02	1.03
414075	U11862	Hs.75741	amiloride binding protein 1 (amine oxide	0.84	1.07
416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	3.67	1.00
417542	J04129	Hs.82269	progesterone-associated endometrial prote	1.28	1.35
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	1.00	1.00
419502	ALJ076704		fibrinogen, A alpha polypeptide	13.05	115.00
419631	AW188117	Hs.303154	popeye protein 3	1.00	13.00
420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	1.00	8.00
421155	H87879	Hs.102267	lysyl oxidase	1.00	15.00
421190	U95031	Hs.102482	mucin 5, subtype B, tracheobronchial	1.17	1.55
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	1.46	1.76
421515	Y11339	Hs.105352	GalNAc alpha-2, 6-sialyltransferase 1,	1.00	3.00
421582	A1910275		tetrol factor 1 (breast cancer, estrogen	1.23	1.00
422026	U80736	Hs.110826	trinucleotide repeat containing 9	1.00	52.00
422095	A1868872	Hs.282804	hypothetical protein FLJ22704	4.37	2.34
422311	AF073515	Hs.114948	cytokine receptor-like factor 1	1.15	1.78
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	1.69	3.17
423472	AF041260	Hs.129057	breast carcinoma amplified sequence 1	48.13	72.00
423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	1.00	50.00
424502	AF242398	Hs.149585	lengsin	1.00	1.00
424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	1.00	59.00
424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	21.35	1.00
424960	BE245380	Hs.153952	5' nucleotidase (CD73)	1.00	1.00
425523	AB007948	Hs.158244	KIAA0479 protein	1.00	35.00
426230	AA367019	Hs.241395	protease, serine, 1 (trypsin 1)	1.00	83.00
427701	AA411101	Hs.243886	nuclear autoantigenic sperm protein (his	7.41	34.00
428585	AB007863	Hs.185140	KIAA0403 protein	1.00	6.00
428758	AA433988	Hs.98502	hypothetical protein FLJ14303	1.08	1.13
429170	NM_001394	Hs.2359	dual specificity phosphatase 4	16.18	105.00
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	1.07	1.00
429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	1.59	1.69
430508	AI015435	Hs.104637	ESTs	4.75	7.27
430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	0.94	1.28
431548	AI834273	Hs.9711	novel protein	5.66	15.00
431566	AF176012	Hs.260720	J domain containing protein 1	49.76	37.00
431986	AA536130	Hs.149018	Novel human gene mapping to chromosome 20	1.19	1.47
432375	BE536069	Hs.2962	S100 calcium-binding protein P	1.65	1.06
432677	NM_004482	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polyp	1.00	48.00
433556	W56321	Hs.111460	calcium/calmodulin-dependent protein kin	1.00	19.00
433819	AW511097	Hs.112765	ESTs	3.71	8.00
434001	AW950905	Hs.3697	serine (or cysteine) proteinase inhibito	29.31	72.00
434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fis, clone L	1.00	84.00
434792	AA649253	Hs.132458	ESTs	8.52	44.00
436217	TS3925	Hs.107	fibrinogen-like 1	57.97	31.00
436749	AA584890	Hs.5302	lectin, galactoside-binding, soluble, 4	1.10	1.41
436972	AA284679	Hs.25640	claudin 3	1.59	1.46
437866	AA156781		metallothionein 1E (functional)	3.62	101.00
437935	AW939591	Hs.5940	nucin 13, epithelial transmembrane	1.60	1.39
438915	AA280174	Hs.285681	Williams-Beuren syndrome chromosome regl	1.00	1.00
439451	AF086270	Hs.278554	heterochromatin-like protein 1	23.28	52.00

439759	AL359055	Hs.67709	Homo sapiens mRNA full length insert cDN	1.00	21.00
441031	AI110684	Hs.7645	fibrinogen, B beta polypeptide	1.41	99.00
441377	BE218239	Hs.202656	ESTs	22.03	1.00
443614	AV655386	Hs.7645	fibrinogen, B beta polypeptide	1.00	16.00
443813	AA876372	Hs.93961	Homo sapiens mRNA; cDNA DKFZp667D095 (fr	1.20	1.99
443991	NM_002250	Hs.10082	potassium intermediate/small conductance	5.71	6.87
444670	H58373	Hs.332938	hypothetical protein MGC5370	1.98	38.00
444931	AV652066	Hs.75113	general transcription factor IIA	1.00	54.00
446102	AW168067	Hs.317694	ESTs	1.00	1.00
446163	AA026880	Hs.25252	Homo sapiens cDNA FLJ13603 fis, clone PL	1.00	36.00
446469	BE094848	Hs.15113	homogentisate 1,2-dioxygenase (homogenti	1.00	11.00
447388	AW630534	Hs.76277	Homo sapiens, clone MGC:9381, mRNA, comp	1.24	1.16
447532	AK000614	Hs.18791	hypothetical protein FLJ20607	1.23	1.63
448243	AW369771	Hs.52620	Integrin, beta 8	15.84	1.00
448844	AI591519	Hs.177164	ESTs	1.00	31.00
449444	AW818436	Hs.23590	solute carrier family 16 (monocarboxylic	1.00	83.00
451807	W52854		hypothetical protein FLJ23293 similar to	1.55	35.00
452689	F33868	Hs.284176	transferrin	1.54	1.44
453392	U23762	Hs.32964	SRY (sex determining region Y)-box 11	1.00	16.00
453464	AI884911	Hs.32989	receptor (calcitonin) activity modifying	1.55	2.45
453735	AI066629	Hs.125073	ESTs	1.01	1.30

TABLE 11B

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
410399	11995_1	BE068889 BE068882 AF044311 AF017256 NM_003087 AF037207 AF010126 AA633976 AA872836 BE298825 BE299889 AJ016464 AI684600 AI936527 AA804675 AA394097 AI139933 AA946606 BE171313 AA722407 AA293803 AI468480 AA056035 AA055968 AW798957 AI637713 AA410737 H49348 AA486472 AA411094 AA235594 AA402624 AA443638 AW452137 AA421708 AW265211 AI493266 AA365132 AW966044 AU076704 T74854 T74860 T72098 T73265 T73873 T69180 T74668 T58786 T60385 T73410 T68781 T67845 T67593 T73952 T67864 T60630 T68367 T68401 T53959 T72360 T72099 T60377 T58961 T71712 T72821 T64738 T74645 T72037 T68688 T72063 T73258 T72826 T64242 T68220 T74673 T71800 T68355 T61227 T62738 T69317 T53850 T64692 T73768 T73962 T73382 T68914 T70975 T73400 T60631 T73277 T73203 T70498 T61409 T58925 NM_000508 M54982 T68301 T73729 T69445 T60424 T67922 T67736 T68716 T67755 T74765 T73819 T58719 T74756 T60477 T74863 T61109 T68329 T58850 T71857 T73425 T53736 T68607 T58898 T64309 T72031 T72079 T64305 T71908 T68107 T71916 T73787 T56035 T64425 T71870 T60476 T61376 T67820 T71895 T41006 T69441 T68170 T74617 T71958 T69440 T61875 R05796 H48353 T71914 T53939 T64121 AA693996 T72525 T67779 T68078 AA011465 AA345376 AV654847 AV654272 AV656001 AI064740 T82897 N33594 AA344542 AW805054 AI207457 T61743 AA025737 H94389 AA382695 AA918409 T68044 S82092 T39959 AI017721 AA312395 AA312919 T40156 H66239 AV852989 H38728 R98521 AV656200 R95790 W03250 W00913 AA344136 AV660126 R97923 AA343596 AW470774 AV651256 N54417 AA812862 AW182929 AI111192 H61463 H72060 AA344503 H38639 AI277511 AV661108 AI207625 T47810 AA235252 T27853 T47778 R95746 H70620 AA701463 AW827166 R98475 C20925 AV657287 T71959 T71313 T73920 T73333 T61618 T69293 T69283 T73931 T72178 T72456 AV645639 AV653476 T72957 T72300 T58906 T71457 T70494 T72956 T70495 T68267 T74407 T85778 AA344726 T27854 T74485 T74101 T73868 T71518 T72304 AA343853 T73909 T68070 T72056 H72149 T73493 T73495 AV645993 R02293 T70475 T64751 AA344441 AA343657 AA345732 AA344328 AI110639 AA344603 AF063513 T64686 T68516 T72223 T60507 T67633 R29500 T72517 R02292 T60599 T69206 T70452 T74577 R29366 T61277 T74914 T60352 R29675 T74843 AV645792 AA344408 T69197 T72057 T69368 T69358 T68258 AV650429 T73341 T61702 T74598 T40095 K02272 T40106 AA343045 AA341908 AA341907 AA342807 AA341964 T53747 T72042 T62764 AI064899 AA343060 T67832 T72440 T71770 T68091 T69108 T72449 T69167 T71289 T68251 AV654844 T64375 AA345234 T67598 AA011414 T68036 H48262 AI207557 T68219 W86031 T69081 T64232 R93196 T62136 AV650539 H67459 T72978 AA344583 T60362 H58121 T95711 T72803 T68055 T71715 R29036 T72793 T69122 T64595 T62868 T69139 T68291 T64652 T67971 T46862 AA693592 AI248502 R29454 T64764 T57001 T73052 T71429 T51176 T58856 AV655414 H90426 AA342489 T73666 T67848 T72512 T53835 T67837 T73317 T74273 T69420 T68245 T74380 T67862 T74474 T56068
421582	2041_1	AI910275 X00474 X52003 X05030 NM_003225 AA314326 AA308400 AA506787 AA314825 AI571948 AA507595 AA614579 AA587613 R83818 AA568312 AA614409 AA307578 AI925552 AW950155 AI910083 M12075 BE074052 AW004668 AA578674 AA582084 BE074053 BE074126 BE074140 AA514776 AA588034 BE074051 BE074068 AW009769 AW050690 AA858276 R55369 AI001051 AW050700 AW750216 AA614539 BE074045 AI307407 AW602303 BE073575 AI202532 AA524242 AI970839 AI909751 BE076078 AI909749 R55292
437866	44433_2	AA156781 AW293839 U52054 AA024963 AA778446 BE073977 AW444904 AW602574 BE164040 BE164012 BE163972 BE163974 BE163992 AA837481 AW468444 BE185091 AW468002 AA687333 AA811830 AA581806 AI866686 AI572124 AA043777 AA040926 D20160 AI536733 AA812489 AW874142 AI471883 W84421 AA156850
451807	6865_1	W52854 AL117600 BE208116 BE208432 BE208239 BE082291 AW953423 AA351619 BE180648 BE140560 W60080 AA865478 N90291 AW406552 AW449519 AA936334 AI806539 AA351618 AW449522 AI827626 AA904788 AA380381 AA886045 AA774409 BE003229 Z41756

TABLE 11C

Pkey: Unique number corresponding to an Eos probeset  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 Nt\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
403329	8518120	Plus	98450-96588
406399	9256288	Minus	63448-63554



TABLE 12A: Genes Distinguishing Squamous Cell Carcinoma from Other Lung Diseases and Normal Lung

Table 12A shows about 72 genes upregulated in squamous cell carcinomas of the lung relative to other lung tumors, non-malignant lung disease, and normal lung. These genes were selected from about 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 12B show the accession numbers for those Pkey's lacking Unigenes/D's for table 12A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 12C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 12A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 Unigenes/D: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	Unigenes/D	Unigene Title	R1	R2
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	132.45	4.00
400666			NM_002425:Homo sapiens matrix metallopro	3.25	3.22
401780			NM_005557:Homo sapiens keratin 16 (foca	26.47	10.50
401781			Target Exon	10.33	4.61
401785			NM_002275:Homo sapiens keratin 15 (KRT1	4.13	2.70
401994			Target Exon	61.84	47.00
402075			ENSP00000251056:Plasma membrane calcium	1.00	1.00
404996			Target Exon	1.00	1.00
407839	AA045144	Hs.161566	ESTs	173.91	108.00
408000	L11690	Hs.620	bulous pemphigoid antigen 1 (230/240kD)	151.17	8.00
408522	AJ541214	Hs.46320	Small proline-rich protein SPRK (human,	1.98	1.24
410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	10.04	1.00
415091	AL044872	Hs.77910	3-hydroxy-3-methylglutaryl-Coenzyme A sy	1.00	30.00
415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	24.30	1.00
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	53.29	51.00
417034	NM_006183	Hs.80962	neurotensin	1.00	1.00
417366	BE185289	Hs.1076	small proline-rich protein 1B (cornitin)	8.97	3.27
418663	AK001100	Hs.41690	desmocollin 3	112.17	19.00
418678	NM_001327	Hs.87225	cancer/testis antigen	1.18	1.10
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	1.00	1.00
420783	AJ659638	Hs.99923	lectin, galactoside-binding, soluble, 7	3.04	1.25
421773	W69233	Hs.112457	ESTs	1.12	1.14
421948	L42583	Hs.334309	keratin 6A	51.83	20.25
421978	AJ243662	Hs.110196	NICE-1 protein	1.01	0.91
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	2.37	1.10
422440	NM_004812	Hs.116724	aldo-keto reductase family 1, member B10	47.53	32.00
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	75.02	1.00
423725	AJ403108	Hs.132127	hypothetical protein LOC57822	4.20	1.00
423738	AB002134	Hs.132195	airway trypsin-like protease	10.14	51.00
424012	AW368377	Hs.137559	tumor protein 63 kDa with strong homolog	233.42	68.00
424046	AF027856	Hs.138202	serine (or cysteine) proteinase inhibito	1.00	1.00
424098	AF077374	Hs.139322	small proline-rich protein 3	137.82	54.00
424834	AK001432	Hs.153408	Homo sapiens cDNA FLJ10570 fis, clone NT	56.19	12.00
425550	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	33.45	1.00
427099	AB032953	Hs.173560	odd Ozten-m homolog 2 (Drosophila, mous	4.24	17.00
427335	AA448542	Hs.251677	G antigen 7B	51.83	4.00
428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN G ANT	1.00	1.00
428645	AA431400	Hs.98729	ESTs, Weakly similar to 2017205A dihydro	1.00	16.00
428748	AW593206	Hs.98785	Ksp37 protein	1.00	87.00
429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	2.01	1.18
429538	BE182592	Hs.11261	small proline-rich protein 2A	4.43	2.90
429903	AL134197	Hs.93597	cyclin-dependent kinase 6, regulatory su	11.80	1.00
430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	12.28	41.00
430890	X54232	Hs.2699	glypican 1	1.58	1.40
431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	60.25	28.00
431846	BE019924	Hs.271580	uropodxin 1B	4.49	2.51
433091	Y12642	Hs.3185	lymphocyte antigen 6 complex, locus D	1.20	1.09
434360	AW015415	Hs.127780	ESTs	40.98	27.00
434880	U02388	Hs.101	cytochrome P450, subfamily IVF, polypept	1.00	1.00
435505	AF200492	Hs.211238	interleukin-1 homolog 1	1.00	38.00
435793	AB037734	Hs.4393	KIAA1313 protein	23.68	42.00
436511	AA721252	Hs.291502	ESTs	16.76	14.00
438403	AA806607	Hs.292206	ESTs	1.00	1.00
439285	AL133916		hypothetical protein FLJ20093	46.23	139.00
439606	W79123	Hs.58561	G protein-coupled receptor 87	33.61	1.00
439670	AF088076	Hs.59507	ESTs, Weakly similar to AC004858 3 U1 sm	1.00	1.00
439706	AW372527	Hs.59761	ESTs, Weakly similar to DAP1_HUMAN DEATH	86.55	11.00
440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma	62.88	147.00
441525	AW241867	Hs.127728	ESTs	1.53	1.42
443162	T49951	Hs.9029	DKFZP434G032 protein	31.11	38.00
444378	R41339	Hs.12569	ESTs	1.00	1.00

446292	AF081497	Hs.279682	Rh type C glycoprotein	1.55	1.26
447078	AW885727	Hs.9914	ESTs	47.24	24.00
447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	28.63	1.00
449003	X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o	1.00	1.00
449101	AA205847	Hs.23016	G protein-coupled receptor	2.58	27.00
450832	AW970602	Hs.105421	ESTs	25.17	36.00
452240	AI591147	Hs.61232	ESTs	13.42	1.00
453317	NM_002277	Hs.41696	keratin, hair, acidic,1	1.19	1.27
453830	AA534296	Hs.20953	ESTs	24.92	25.00
454098	WZ7953	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	1.26	1.11
455601	AI368680	Hs.816	SRY (sex determining region Y)-box 2	206.11	1.00

TABLE 12B

Pkey: Unique Eos probeset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Pkey	CAT Number	Accession
439285	47055_1	AL133916 N79113 AF088101 N76721 AW950828 AA364013 AW955684 AI346341 AI867454 N54784 AI655270 AI421279 AW014882 AA775552 N62351 N59253 AA626243 AI341407 BE175639 AA456968 AI358918 AA457077

TABLE 12C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
400666	8118496	Plus	17982-18115,20297-20456
401780	7249190	Minus	28397-28617,28920-29045,29135-29296,29411-29567,29705-29787,30224-30573
401781	7249190	Minus	83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,85290-86814
401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
401994	4153858	Minus	42904-43124,43211-43336,44607-44763,45199-45281,46337-46732
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
404996	6007890	Plus	37999-38145,38652-38998,39727-39872,40557-40674,42351-42450

TABLE 13A: Genes Distinguishing Non-Malignant Lung Disease from Lung Tumors and Normal lung

Table 13A shows about 23 genes upregulated in non-malignant lung disease relative to lung tumors and normal lung. These genes were selected from about 59680 probesets on the Eos/Alfymetrix Hu03 Genechip array.

Table 13B show the accession numbers for those Pkey's lacking UnigeneID's for table 13A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 13C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 13A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigeneID: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Average of lung tumors (including squamous cell carcinomas, adenocarcinomas, small cell carcinomas, granulomatous and carcinoid tumors) divided by the average of normal lung samples  
 R2: Average of non-malignant lung disease samples (including bronchitis, emphysema, fibrosis, atelectasis, asthma) divided by the average of normal lung samples

Pkey	ExAccn	UnigeneID	Unigene Title	R1	R2
408562	AA36323	Hs.31141	Homo sapiens mRNA for KIAA1568 protein,	1.00	230.00
409031	AA376836	Hs.76728	ESTs	1.00	128.00
412372	R65998	Hs.285243	hypothetical protein FLJ22029	1.00	173.00
415910	U20350	Hs.78913	chemokine (C-X3-C) receptor 1	1.00	145.00
417511	AL049176	Hs.82223	chordin-like	1.00	179.00
418819	AA228776	Hs.191721	ESTs	1.00	140.00
422060	R20893	Hs.325823	ESTs, Moderately similar to ALUS_HUMAN A	1.00	156.00
424585	AA464840	Hs.131987	ESTs	1.00	167.00
426753	T69832	Hs.170278	ESTs	1.00	141.00
429496	AA453800	Hs.192793	ESTs	1.00	138.00
430719	AA488988	Hs.293796	ESTs	1.00	133.00
431089	BE041395		ESTs, Weakly similar to unknown protein	23.32	941.00
431385	BE178536	Hs.11090	membrane-spanning 4-domains, subfamily A	1.00	157.00
431728	NM_007351	Hs.268107	multimerin	1.00	157.00
436532	AA721522		gbzv54h12.r1 NCL_CGAP_Ew1 Homo sapiens	1.00	218.00
437960	AI669586	Hs.222194	ESTs	1.00	147.00
438202	AW169287	Hs.22588	ESTs	1.00	141.00
441499	AW298235	Hs.101689	ESTs	1.00	167.00
444513	AL120214	Hs.7117	glutamate receptor, ionotropic, AMPA 1	1.00	151.00
448253	H25899	Hs.201591	ESTs	1.00	141.00
453636	R67837	Hs.169872	ESTs	1.00	116.00
458332	AI000341	Hs.220491	ESTs	1.00	192.00
459587	AA031956		gbzkl5e04.s1 Soares_pregnant_uterus_NbH	1.00	154.00

TABLE 13B

Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
431089	327825_1	BE041395 AA491826 AA621946 AA715980 AA666102
436532	421802_1	AA721522 AW975443 T93070

TABLE 13C

Pkey: Unique number corresponding to an Eos probeset  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076

TABLE 14A: Preferred Utility and Subcellular Localization for Potential Lung Disease Targets

Table 14A shows the subcellular localization and preferred utility for the genes appearing in Tables 9A and 10A. mAb symbolizes monoclonal antibody, diag symbolizes diagnostic, s.m. symbolizes small molecule, and CTL symbolizes cytotoxic lymphocytic ligand. These genes were selected from 59680 probesets on the Eos/Affymetrix Hu03 Genechip array.

Table 14B show the accession numbers for those Pkey's lacking UnigenelD's for table 14A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 14C show the genomic positioning for those Pkey's lacking UnigenelD's and accession numbers in table 14A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique Eos probeset identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UnigenelD: Unigenel number

Unigenel Title: Unigenel gene title

Pref.Utility: Preferred Utility

Pred.Loc: Predicted subcellular localization

Pkey	ExAccn	UnigenelD	Unigenel Title	Pref Utility	Pred. Loc
400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	mAb & diag & s.m.	extracellular
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	mAb	plasma membrane
402075			ENSP00000251056**Plasma membrane calcium	mAb & diag	secreted
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagonist	diag	secreted
408243	Y00787	Hs.824	interleukin 8	diag	secreted
408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	mAb & s.m.	plasma membrane
408908	BE296227	Hs.250822	serine/threonine kinase 15	s.m.	cytoplasm
409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	CTL & diag	secreted
409103	AF251237	Hs.112208	XAGE-1 protein	CTL	nuclear
409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini	diag	secreted
409832	W74001	Hs.55279	serine (or cysteine) proteinase inhibitor	diag	secreted
409757	NM_001898	Hs.123114	cystatin SN	diag	extracellular
409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.	CTL	nuclear
409956	AW103364	Hs.727	inhibin, beta A (activin A, activin AB a	diag	extracellular
410001	AB041036	Hs.57771	kallikrein 11	diag	extracellular
410407	X66839	Hs.63287	carbonic anhydrase IX	mAb & s.m.	plasma membrane
410418	D31382	Hs.63325	transmembrane protease, serine 4	mAb & diag & s.m.	plasma membrane
412140	AA219691	Hs.73625	RAB6 interacting, kinesin-like (rakines	s.m.	
412719	AW016610	Hs.816	ESTs	s.m.	nuclear
414774	X02419	Hs.77274	plasminogen activator, urokinase	diag	extracellular
414883	AA926960		CDC28 protein kinase 1	s.m.	
415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	CTL & diag	extracellular
415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor	mAb & diag & s.m.	secreted
415817	U88967	Hs.78887	protein tyrosine phosphatase, receptor-t	mAb & s.m.	plasma membrane
416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara	diag	extracellular
417034	NM_006183	Hs.80952	neurotensin	diag	extracellular
417079	U65590	Hs.81134	interleukin 1 receptor antagonist	diag	extracellular
417308	H60720	Hs.81892	KIAA0101 gene product	s.m.	mitochondrial
417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor	mAb & diag	secreted
417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein	mAb	plasma membrane
417933	X02308	Hs.82962	thymidylate synthetase	s.m.	endoplasmic reticulum
418478	U88945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	s.m.	cytoplasm
418506	AA084248	Hs.85339	G protein-coupled receptor 39	mAb & s.m.	plasma membrane
418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)	CTL	cytoplasmic
419121	AA374372	Hs.89626	parathyroid hormone-like hormone	diag	secreted
419171	NM_002846	Hs.89655	protein tyrosine phosphatase, receptor t	mAb & s.m.	plasma membrane
419183	U60669	Hs.89663	cytochrome P450, subfamily XXIV (vitamin	CTL & s.m.	mitochondrial
419216	AJ076718	Hs.164021	small inducible cytokine subfamily B (Cy	diag	secreted
419235	AW470411	Hs.288433	neurotrophin	mAb & diag	plasma membrane
419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	mAb & s.m.	plasma membrane
419556	U29615	Hs.91093	chitinase 1 (chitinotrioidase)	mAb & diag	extracellular*
420610	AJ683183	Hs.99348	distal-less homeo box 5	CTL	nuclear
421110	AJ250717	Hs.1355	cathepsin E	sm & diag	extracellular
421379	Y15221	Hs.103962	small inducible cytokine subfamily B (Cy	diag	secreted
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	mAb & s.m.	plasma membrane
421552	AF026692	Hs.105700	secreted frizzled-related protein 4	diag	secreted
421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	mAb & s.m.	plasma membrane
421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR	mAb & s.m.	plasma membrane
422109	ST3265	Hs.1473	gastrin-releasing peptide	diag	secreted
422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	diag	secreted
422282	AF019225	Hs.114309	apolipoprotein L	diag	secreted
422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	s.m.	nuclear
422424	AJ186431	Hs.296638	prostate differentiation factor	diag	extracellular
422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur	s.m.	cytoplasm
422809	AK001379	Hs.121028	hypothetical protein FLJ10549	s.m.	nuclear
422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse	diag	extracellular
422956	BE545072	Hs.122579	ECT2 protein (Epithelial cell transformi	CTL & s.m.	
423634	AW959908	Hs.1690	heparin-binding growth factor binding pr	diag	
423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	mAb & diag & s.m.	secreted
423961	D13666	Hs.136348	perostin (OSF-2os)	mAb & diag	extracellular
424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor	diag	secreted
424381	AA285249	Hs.146329	protein kinase Ctk2	s.m.	nuclear

	424502	AF242388	Hs.149585	lengsin	s.m.	cytoplasmic
	424503	NM_002205	Hs.149809	integrin, alpha 5 (fibronectin receptor,	mAb & s.m.	plasma membrane
	424507	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	diag	extracellular
5	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	mAb & diag & s.m.	secreted
	425322	U83630	Hs.155637	protein kinase, DNA-activated, catalytic	s.m.	cytoplasmic
	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	mAb	plasma membrane
	425734	AF056209	Hs.159396	peptidylglycine alpha-oxidizing monooxyg	s.m.	
	425776	U25128	Hs.159499	parathyroid hormone receptor 2	mAb & diag	plasma membrane
10	425852	AF1041504	Hs.159651	death receptor 6, TNF superfamily member	mAb & s.m.	plasma membrane
	426215	AW963419	Hs.155223	stanniocalcin 2	mAb & diag	secreted
	426427	M86699	Hs.169840	TTK protein kinase	CTL & s.m.	nuclear
	426514	BE616633	Hs.170195	bone morphogenetic protein 7 (osteogenic	mAb & diag	secreted
	427335	AA448542	Hs.251677	G antigen 7B	CTL	cytoplasmic
	427747	AW411425	Hs.180655	serine/threonine kinase 12	s.m.	cytoplasmic
15	428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	diag	
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	mAb & diag & s.m.	extracellular
	428450	NM_014791	Hs.184339	KIAA0175 gene product	s.m.	nuclear
	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to	s.m.	nuclear
20	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino	mAb & s.m.	plasma membrane
	428664	AK001666	Hs.189095	similar to SALL1 (sal (Drosophila)-like	CTL & s.m.	nuclear
	428698	AA852773	Hs.334838	KIAA1866 protein	mAb	
	428748	AW593206	Hs.98785	Kap37 protein	diag	extracellular
	428758	AA433988	Hs.98502	CA125 antigen; mucin 16	diag	mitochondria*
25	428969	AF120274	Hs.194689	artemin	diag	extracellular
	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	mAb & s.m.	plasma membrane
	429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	mAb & s.m.	plasma membrane
	429547	AW009166	Hs.99376	ESTs	diag	secreted
	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	mAb & diag	secreted
30	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	s.m.	
	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	mAb & s.m.	plasma membrane
	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	diag	extracellular
	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha	mAb & s.m.	plasma membrane
	431846	BE019924	Hs.271580	uroplakin 1B	mAb & diag	plasma membrane
35	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	mAb & diag	plasma membrane
	432201	AF538613	Hs.298241	Transmembrane protease, serine 3	mAb & diag & s.m.	plasma membrane
	433001	AF217513	Hs.279905	clone HQ0310 PRO0310p1	s.m.	nuclear
	435505	AF200492	Hs.211238	Interleukin-1 homolog 1	diag	secreted
40	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con	s.m.	
	437016	AL076916	Hs.5398	guanine monophosphate synthetase	s.m.	cytoplasm
	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an	CTL	ER
	437789	AF581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheri	CTL	nuclear
	437852	BE001836	Hs.256897	ESTs, Weakly similar to J365012.1 [H.se	mAb & s.m.	plasma membrane
	439223	AW238299	Hs.250618	UL16 binding protein 2	mAb	plasma membrane
45	439477	W69813	Hs.58042	ESTs, Moderately similar to GFR3_HUMAN G	mAb & s.m.	
	439606	W79123	Hs.58561	G protein-coupled receptor 87	mAb & s.m.	plasma membrane
	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	mAb & s.m.	plasma membrane
	440006	AK000517	Hs.6844	NALP2 protein; PYRIN-Containing APAF1-li	s.m.	nuclear
	441362	BE514410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Ra	s.m.	
50	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447	mAb & s.m.	plasma membrane
	443247	BE614387	Hs.333893	c-Myc target JPO1	CTL	extracellular*
	443426	AF098158	Hs.9329	chromosome 20 open reading frame 1	CTL	
	443859	NM_013409	Hs.9914	folistatin	diag	extracellular
	444006	BE395085	Hs.10086	type I transmembrane protein Fn14	mAb	plasma membrane
55	444371	BE540274	Hs.239	forkhead box M1	s.m.	nuclear
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheri	diag	secreted
	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	mAb & diag	plasma membrane
	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6	mAb & diag	secreted
	446619	AL076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	diag	secreted
60	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	diag	extracellular
	447033	AF357412	Hs.157601	ESTs	CTL & diag	secreted
	447342	AF199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	CTL	
	448243	AW369771	Hs.52620	Integrin, beta 8	mAb & s.m.	plasma membrane
	448844	AF581519	Hs.177164	ESTs	mAb & s.m.	
65	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	mAb	plasma membrane
	449722	BE280074	Hs.23980	cyclin B1	s.m.	cytoplasm
	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte	mAb & s.m.	plasma membrane
	450375	AA009647		a disintegrin and metalloproteinase doma	mAb & diag & s.m.	plasma membrane
	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	mAb & diag	plasma membrane
70	450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	diag	secreted
	451668	Z43948	Hs.326444	cartilage acidic protein 1	mAb & diag	plasma membrane
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 f1s, clone PL	diag	
	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	diag	extracellular
	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	mAb	plasma membrane
75	452838	U65011	Hs.30743	preferentially expressed antigen in mela	CTL	nuclear
	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso	CTL & s.m.	nuclear
	457469	AF693815	Hs.127179	cryptic gene	diag	secreted

TABLE 14B

80 Pkey: Unique Eos probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey CAT Number Accession

WO 02/086443

PCT/US02/12476

414883 15024\_1 AA926960 AA926959 W76521 W24270 W21526 AA037172 BE267636 H83186 AA469909 N86396 AA001348 BE535736 AA081745 BE566245  
AA082436 H72525 H77575 N49786 W80565 H78746 BE569085 W04339 R96127 T55938 BE279271 AW960304 T29812 AA476873 BE297387  
AA292753 AA177048 NM\_001826 X54941 BE314366 AA908783 A1719075 BE270172 BE269819 AA889955 A1204630 W25243 A1935150  
AA872039 W72395 T99630 A1422691 H98460 N31428 BE255916 H03265 A1857576 AA776920 AA910544 AA459522 AA293140 AW514667  
R75953 AW662396 AA662522 A1865147 A1423153 AW262230 AA584410 AA583187 AW024595 AW069734 A1828996 AA282997 AA876046  
AW613002 AA527373 AW972459 A1831360 AA621337 AA100926 AA772418 AA594628 A1033892 W95096 A1034317 AA398727 A1085031  
N95210 A1459432 A1041437 AA932124 AA627684 AA935829 A1004827 A1423513 A1094597 H42079 R54703 A1630359 AA617681 AA978045  
AA643280 W44561 A1891988 A1537692 A1090262 AA740817 A1312104 A1911822 AA416871 A1185409 AA129784 AA701623 A1075239  
A1139549 AA633548 A1339996 A1336880 AA399239 A1078708 A1085351 A1362835 A1346618 A1146955 A1989380 A1348243 N92892 AA765850  
A1494230 A1278887 AA962596 A1492600 W80435 AA001979 R97424 A1129015 N24127 AA157451 AA235549 AA459292 AA037114 AA129785  
A1494211 AW059601 AW886710 R92790 N59755 A1361128 AW589407 H47725 H97534 H48076 H48450 T89631 AW300758 H03431 R76789  
AA954344 H77576 R96823 A1457100 N92845 N49682 H42038 BE220698 BE220715 H99552 AA701624 N74173 R54704 H79520 H72923  
H03266 BE261919 AA769633 AA480310 AA507454 AA910586 A1203723 AW104725 W25611 W25071 T88980 H03513 T77589 R99156  
W95095 R97470 AA702275 T77551 AA911952 H82956 N83673 AA283672  
450375 83327\_1 AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157467 BE157601 H04384 W46291 AW663674 H04021 H01532  
AA190993 H03231 H59605 H01642 AA852876 AA113758 AA628915 AA746952 A1161014 AA099554 R69067

TABLE 14C

Pkey: Unique number corresponding to an Eos probeset  
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA  
sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
Strand: Indicates DNA strand from which exons were predicted.  
NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035, 122804-122921, 124019-124161, 124455-124610, 125672-126076

TABLE 15A: Information for all sequences in Table 16

Table 15A shows the Seq ID No, Pkey, ExAccn, UnigenelD, and Unigene Title for all of the sequences in Table 16.

5 Table 15B show the accession numbers for those Pkey's lacking UnigenelD's for table 15A. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10 Table 15C show the genomic positioning for those Pkey's lacking Unigene ID's and accession numbers in table 15A. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

15 Seq ID No: Sequence ID number  
Pkey: Unique Eos probeset identifier number  
ExAccn: Exemplar Accession number, Genbank accession number  
UnigenelD: Unigene number  
Unigene Title: Unigene gene title

20	Seq ID No:	Pkey	ExAccn	UnigenelD	Unigene Title
	Seq ID No: 1 & 2	410407	X66839	Hs.63287	carbonic anhydrase IX
	Seq ID No: 3 & 4	412719	AW016610	Hs.816	ESTs
	Seq ID No: 5 & 6	417034	NM_006183	Hs.80962	neurotensin
25	Seq ID No: 7 & 8	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
	Seq ID No: 9 & 10	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 11 & 12	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 13 & 14	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
	Seq ID No: 15 & 16	407788	BE514982	Hs.38991	S100 calcium-binding protein A2
30	Seq ID No: 17 & 18	439285	AL133916		hypothetical protein FLJ20093
	Seq ID No: 19 & 20	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kalinin
	Seq ID No: 21 & 22	120486	AW368377	Hs.137669	tumor protein 63 kDa with strong homolog
	Seq ID No: 23 & 24	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	Seq ID No: 25 & 26	412140	AA219891	Hs.73625	RAB6 interacting, kinesin-like (raklines
35	Seq ID No: 27 & 28	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
	Seq ID No: 29 & 30	452838	U65011	Hs.30743	preferentially expressed antigen in mela
	Seq ID No: 31 & 32	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 33 & 34	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 35 & 36	409632	W74001	Hs.55279	serine (or cysteine) proteinase inhibitor
40	Seq ID No: 37 & 38	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
	Seq ID No: 39 & 40	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
	Seq ID No: 41 & 42	431846	BE019924	Hs.271580	uropod 1B
	Seq ID No: 43 & 44	418830	BE513731	Hs.88959	hypothetical protein MGC4816
	Seq ID No: 45 & 46	424098	AF077374	Hs.139322	small proline-rich protein 3
45	Seq ID No: 47 & 48	443648	AI085377	Hs.143610	ESTs
	Seq ID No: 49	311034	BE567130	Hs.311369	ESTs, Highly similar to NKGD_HUMAN NKG2-
	Seq ID No: 50 & 51	408522	AI541214	Hs.46320	Small proline-rich protein SPRK (human,
	Seq ID No: 52 & 53	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 54 & 55	435505	AF200492	Hs.211238	interleukin-1 homolog 1
50	Seq ID No: 56 & 57	417366	BE185289	Hs.1076	small proline-rich protein 18 (comifin)
	Seq ID No: 58 & 59	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
	Seq ID No: 60 & 61	441020	W79283	Hs.35962	ESTs
	Seq ID No: 62 & 63	423217	NM_000094	Hs.1640	collagen, type VII, alpha 1 (epidermolys
	Seq ID No: 64 & 65	429538	BE182592	Hs.11261	small proline-rich protein 2A
55	Seq ID No: 66 & 67	448733	NM_005629	Hs.187958	solute carrier family 6 (neurotransmitte
	Seq ID No: 68 & 69	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 70 & 71	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 72 & 73	444371	BE540274	Hs.239	forkhead box M1
	Seq ID No: 74 & 75	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
60	Seq ID No: 76 & 77	422168	AA586894	Hs.112408	S100 calcium-binding protein A7 (psorias
	Seq ID No: 78 & 79	429259	AA420450	Hs.292911	Plakophilin
	Seq ID No: 80 & 81	426440	BE382756	Hs.169902	solute carrier family 2 (facilitated glu
	Seq ID No: 82 & 83	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
	Seq ID No: 84 & 85	423662	AK001035	Hs.130881	B-cell CLL/lymphoma 11A (zinc finger pro
65	Seq ID No: 86 & 87	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
	Seq ID No: 88 & 89	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
	Seq ID No: 90 & 91	417389	BE260964	Hs.82045	midkine (neurile growth-promoting factor
	Seq ID No: 92 & 93	423634	AW959908	Hs.1690	heparin-binding growth factor binding pr
	Seq ID No: 94 & 95	417515	L24203	Hs.82237	ataxia-telangiectasia group D-associated
70	Seq ID No: 96 & 97	441362	BE614410	Hs.23044	RAD51 (S. cerevisiae) homolog (E coli Re
	Seq ID No: 98 & 99	425322	U63630	Hs.155637	protein kinase, DNA-activated, catalytic
	Seq ID No: 100 & 101		X76342	Hs.389	alcohol dehydrogenase 7 (class IV), mu o
	Seq ID No: 102 & 103	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
	Seq ID No: 104 & 105	409103	AF251237	Hs.112208	XAGE-1 protein
75	Seq ID No: 106 & 107	417542	J04129	Hs.82269	progestagen-associated endometrial prote
	Seq ID No: 108 & 109	428471	X57348	Hs.184510	stratelin
	Seq ID No: 110 & 111	418004	U37519	Hs.87539	aldehyde dehydrogenase 3 family, member
	Seq ID No: 112 & 113	414751	AU077228	Hs.77256	enhancer of zeste (Drosophila) homolog 2
	Seq ID No: 114 & 115	418203	X54942	Hs.83758	CDC28 protein kinase 2
80	Seq ID No: 116	447343	AA256641	Hs.236894	ESTs, Highly similar to S02392 alpha-2-m
	Seq ID No: 117 & 118	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 119 & 120	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
	Seq ID No: 121 & 122	446989	AK001898	Hs.16740	hypothetical protein FLJ11036
	Seq ID No: 123 & 124	457819	AA057484	Hs.35406	ESTs, Highly similar to unnamed protein
85	Seq ID No: 125 & 126	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B

	Seq ID No: 127 & 128	414430	AI346201	Hs.76118	ubiquitin carboxyl-terminal esterase L1
	Seq ID No: 129 & 130	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 131 & 132	100568	L05424	Hs.169610	CD44 antigen (homing function and Indian
5	Seq ID No: 133 & 134	458933	AI638429	Hs.24763	RAN binding protein 1
	Seq ID No: 135 & 136	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 137 & 138	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 139 & 140	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
	Seq ID No: 141 & 142	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me
10	Seq ID No: 143 & 144	446269	AW263155	Hs.14559	hypothetical protein FLJ10540
	Seq ID No: 145 & 146	422765	AW409701	Hs.1578	baculoviral IAP repeat-containing 5 (sur
	Seq ID No: 147 & 148	436481	AA379597	Hs.5199	HSPC150 protein similar to ubiquitin-con
	Seq ID No: 149 & 150	440325	NM_003812	Hs.7164	a disintegrin and metalloproteinase doma
	Seq ID No: 151 & 152	439606	W79123	Hs.58561	G protein-coupled receptor 87
15	Seq ID No: 153 & 154	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 155 & 156	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 157 & 158	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 159 & 160	453884	AA355925	Hs.36232	KIAA0186 gene product
	Seq ID No: 161 & 162	404877			NM_005365: Homo sapiens melanoma antigen,
20	Seq ID No: 163 & 164	413129	AF292100	Hs.104613	RP42 homolog
	Seq ID No: 165 & 166	413281	AA861271	Hs.222024	transcription factor BMAL2
	Seq ID No: 167 & 168	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 169 & 170	416819	U77735	Hs.80205	pim-2 oncogene
	Seq ID No: 171 & 172	451320	AW118072		diacylglycerol kinase, zeta (104kD)
25	Seq ID No: 173 & 174	418543	NM_005329	Hs.85982	hyaluronan synthase 3
	Seq ID No: 175 & 176	454034	NM_000691	Hs.575	aldehyde dehydrogenase 3 family, member
	Seq ID No: 177 & 178	426397	J04088	Hs.126346	topoisomerase (DNA) II alpha (170kD)
	Seq ID No: 179 & 180	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 181 & 182	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
30	Seq ID No: 183 & 184	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 185 & 186	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 187 & 188	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 189 & 190	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 191 & 192	448993	AI471630	Hs.8127	KIAA0144 gene product
35	Seq ID No: 193 & 194	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 195 & 196	430393	BE185030	Hs.241305	estrogen-responsive B box protein
	Seq ID No: 197 & 198	425057	AA826434	Hs.1619	achaete-scute complex (Drosophila) homol
	Seq ID No: 199 & 200	420462	AF050147	Hs.97932	chondromodulin 1 precursor
	Seq ID No: 201 & 202	102963	X02404	Hs.274534	calcitonin-related polypeptide, beta
40	Seq ID No: 203 & 204	100576	X00356	Hs.37058	calcitonin/calcitonin-related polypeptid
	Seq ID No: 205 & 206	101175	U82671	Hs.36980	melanoma antigen, family A, 2
	Seq ID No: 207 & 208	429038	AL023513	Hs.194768	seizure related gene 6 (mouse)-like
	Seq ID No: 209 & 210	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 211 & 212	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
45	Seq ID No: 213 & 214	131827	AJ003112	Hs.34780	doublecortin; fissencephaly, X-linked (d
	Seq ID No: 215 & 216	428182	BE386042	Hs.293317	ESTs, Weakly similar to GGC1_HUMAN GANT
	Seq ID No: 217 & 218	427335	AA448542	Hs.251677	G antigen 7B
	Seq ID No: 219 & 220	409420	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalini
	Seq ID No: 221 & 222	114346	AL137256	Hs.130489	ATPase, aminophospholipid transporter-II
50	Seq ID No: 223 & 224	438956	W00847	Hs.135056	Human DNA sequence from clone RP5-850E9
	Seq ID No: 225 & 226	404440			NM_021048: Homo sapiens melanoma antigen,
	Seq ID No: 227 & 228	415669	NM_005025	Hs.78589	serine (or cysteine) proteinase inhibitor
	Seq ID No: 229 & 230	103312	Y12642	Hs.3185	lysosomal
	Seq ID No: 231 & 232	320843	BE069288	Hs.34744	Homo sapiens mRNA; cDNA DKFZp547C136 (tr
55	Seq ID No: 233	429065	AF53247	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT
	Seq ID No: 234 & 235	446102	AW168067	Hs.317694	ESTs
	Seq ID No: 236 & 237	330495	U47924	Hs.71642	guanine nucleotide binding protein (G pr
	Seq ID No: 238	413573	AI733869	Hs.149089	ESTs
	Seq ID No: 239 & 240	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
60	Seq ID No: 241 & 242	428479	Y00272	Hs.334562	cell division cycle 2, G1 to S and G2 to
	Seq ID No: 243 & 244	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 245	437915	AI637993	Hs.202312	Homo sapiens clone N11 NTERA2D1 teratoca
	Seq ID No: 246 & 247	441553	AA281219	Hs.121296	ESTs
	Seq ID No: 248 & 249	331692	AI683487	Hs.152213	wingless-type MMTV integration site fami
65	Seq ID No: 250 & 251	429413	NM_014058	Hs.201877	DESC1 protein
	Seq ID No: 252 & 253	422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis
	Seq ID No: 254 & 255	448357	N20169	Hs.108923	RAB38, member RAS oncogene family
	Seq ID No: 256 & 257	446292	AF081497	Hs.279682	Rh type C glycoprotein
	Seq ID No: 258 & 259	416209	AA236776	Hs.79078	MAD2 (mitotic arrest deficient, yeast, h
70	Seq ID No: 260 & 261	453922	AF063306	Hs.36708	budding uninhibited by benzimidazoles 1
	Seq ID No: 262 & 263	424046	AF027866	Hs.138202	serine (or cysteine) proteinase inhibitor
	Seq ID No: 264 & 265	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 266 & 267	429228	AI553633	Hs.326447	ESTs
	Seq ID No: 268 & 269	409757	NM_001898	Hs.123114	cystatin SN
75	Seq ID No: 270 & 271	411089	AA456454	Hs.214291	cell division cycle 2-like 1 (PITSLRE pr
	Seq ID No: 272 & 273	436511	AA721252	Hs.291502	ESTs
	Seq ID No: 274 & 275	428969	AF120274	Hs.194689	arlemmin
	Seq ID No: 276 & 277	428969	AF120274	Hs.194689	arlemmin
	Seq ID No: 278 & 279	428969	AF120274	Hs.194689	arlemmin
80	Seq ID No: 280 & 281	428969	AF120274	Hs.194689	arlemmin
	Seq ID No: 282	407137	T97307		gbye53h05.s1 Soares fetal liver spleen
	Seq ID No: 283 & 284	412723	AA648459	Hs.335951	hypothetical protein AF301222
	Seq ID No: 285 & 286	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 287 & 288	405770			NM_002362: Homo sapiens melanoma antigen,
85	Seq ID No: 289 & 290	439453	BE264974	Hs.65666	thyroid hormone receptor interactor 13
	Seq ID No: 291 & 292	414774	X02419	Hs.77274	plasminogen activator, urokinase



5	Seq ID No: 293 & 294	424629	M90656	Hs.151393	glutamate-cysteine ligase, catalytic sub
	Seq ID No: 295 & 296	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 297 & 298	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 299 & 300	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 301 & 302	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
10	Seq ID No: 303 & 304	437789	AI581344	Hs.127812	ESTs, Weakly similar to T17330 hypotheti
	Seq ID No: 305 & 306	453968	AA847843	Hs.62711	High mobility group (nonhistone chromoso
	Seq ID No: 307 & 308	403478			NM_022342:Homo sapiens kinesin protein 9
	Seq ID No: 309	441525	AW241867	Hs.127728	ESTs
	Seq ID No: 310 & 311	434105	AW952124	Hs.13094	presentilins associated rhomboid-like pro
15	Seq ID No: 312 & 313	428810	AF068236	Hs.193788	nitric oxide synthase 2A (inducible, hep
	Seq ID No: 314 & 315	413691	AB023173	Hs.75478	ATPase, Class VI, type 11B
	Seq ID No: 316 & 317	423934	U89995	Hs.159234	forkhead box E1 (thyroid transcription f
	Seq ID No: 318 & 319	409228	R16811	Hs.22010	ESTs, Weakly similar to 2109260A B cell
	Seq ID No: 320 & 321	425734	AF056209	Hs.159396	peptidylglycine alpha-amidating monooxyg
20	Seq ID No: 322 & 323	413582	AW295647	Hs.71331	hypothetical protein MGC5350
	Seq ID No: 324 & 325	438403	AA806607	Hs.292206	ESTs
	Seq ID No: 326 & 327	403329			unnamed protein product [Homo sapiens]
	Seq ID No: 328 & 329	409893	AW247090	Hs.57101	minichromosome maintenance deficient (S.
	Seq ID No: 330 & 331	119073	BE245360	Hs.279477	v-ets erythroblastosis virus E26 oncogen
25	Seq ID No: 332 & 333	113195	H83265	Hs.8881	ESTs, Weakly similar to S41044 chromosom
	Seq ID No: 334 & 335	102283	AW161552	Hs.83381	guanine nucleotide binding protein 11
	Seq ID No: 336 & 337	101345	NM_005795	Hs.152175	calcitonin receptor-like
	Seq ID No: 338 & 339	103280	U84722	Hs.76206	cadherin 5, type 2, VE-cadherin (vascula
	Seq ID No: 340 & 341	102012	BE259035	Hs.118400	singed (Drosophila)-like (sea urchin fas
30	Seq ID No: 342 & 343	105729	H46612	Hs.293815	Homo sapiens HSPC285 mRNA, partial cds
	Seq ID No: 344 & 345	134299	AW580939	Hs.97199	complement component C1q receptor
	Seq ID No: 346 & 347	412719	AW016610	Hs.816	ESTs
	Seq ID No: 348 & 349	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
	Seq ID No: 350 & 351	128924	BE279383	Hs.26557	plakophilin 3
35	Seq ID No: 352 & 353	100486	T19006	Hs.10842	RAN, member RAS oncogene family
	Seq ID No: 354 & 355	419121	AA374372	Hs.89626	parathyroid hormone-like hormone
	Seq ID No: 356 & 357	409459	D86407	Hs.54481	low density lipoprotein receptor-related
	Seq ID No: 358 & 359	330493	M27826		endogenous retroviral protease
	Seq ID No: 360 & 361	417666	AW067903	Hs.82772	collagen, type XI, alpha 1
40	Seq ID No: 362 & 363	418113	AI272141	Hs.83484	SRY (sex determining region Y)-box 4
	Seq ID No: 364 & 365	437016	AU076916	Hs.5398	guanine monophosphate synthetase
	Seq ID No: 366 & 367	429612	AF062649	Hs.252587	pituitary tumor-transforming 1
	Seq ID No: 368 & 369	440704	M69241	Hs.182	insulin-like growth factor binding prote
	Seq ID No: 370 & 371	431221	AA449015	Hs.286145	SRB7 (suppressor of RNA polymerase B, ye
45	Seq ID No: 372 & 373	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 374 & 375	431565	AF161470	Hs.260622	butyrate-induced transcript 1
	Seq ID No: 376 & 377	132354	BE185289	Hs.1076	small proline-rich protein 1B (comifin)
	Seq ID No: 378 & 379	424441	X14850	Hs.147097	H2A histone family, member X
	Seq ID No: 380 & 381	103768	AF086009	Hs.296398	gtx:Homo sapiens full length insert cDNA
50	Seq ID No: 382 & 383	417512	X76534	Hs.82226	glycoprotein (transmembrane) nmb
	Seq ID No: 384 & 385	425266	J00077	Hs.155421	alpha-fetoprotein
	Seq ID No: 386 & 387	424503	NM_002205	Hs.149609	Integrin, alpha 5 (fibronectin receptor,
	Seq ID No: 388 & 389	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin
	Seq ID No: 390 & 391	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
55	Seq ID No: 392 & 393	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial
	Seq ID No: 394 & 395	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino
	Seq ID No: 396 & 397	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2
	Seq ID No: 398 & 399	418506	AA084248	Hs.85339	G protein-coupled receptor 39
	Seq ID No: 400 & 401	423851	D13666	Hs.136348	periostin (OSF-2os)
60	Seq ID No: 402 & 403	414812	X72755	Hs.77357	monokine induced by gamma interferon
	Seq ID No: 404 & 405	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	Seq ID No: 406 & 407	417433	BE270266	Hs.82128	5T4 oncofetal trophoblast glycoprotein
	Seq ID No: 408 & 409	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (pse
	Seq ID No: 410 & 411	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy
65	Seq ID No: 412 & 413	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti
	Seq ID No: 414 & 415	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated
	Seq ID No: 416 & 417	411789	AF245505	Hs.72157	Adican
	Seq ID No: 418 & 419	428698	AA852773	Hs.334838	KIAA1866 protein
	Seq ID No: 420 & 421	450098	W27249	Hs.8109	hypothetical protein FLJ21080
70	Seq ID No: 422 & 423	421552	AF026692	Hs.105700	secreted frizzled-related protein 4
	Seq ID No: 424 & 425	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR
	Seq ID No: 426 & 427	450375	AA009647		a disintegrin and metalloproteinase doma
	Seq ID No: 428 & 429	426215	AW963419	Hs.155223	stanniocalcin 2
	Seq ID No: 430 & 431	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin
75	Seq ID No: 432 & 433	432201	AI538613	Hs.298241	Transmembrane protease, serine 3
	Seq ID No: 434 & 435	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph
	Seq ID No: 436 & 437	442117	AW664964	Hs.128899	ESTs; hypothetical protein for IMAGE:447
	Seq ID No: 438 & 439	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn
	Seq ID No: 440 & 441	447033	AI357412	Hs.157601	ESTs
80	Seq ID No: 442 & 443	447033	AI357412	Hs.157601	ESTs
	Seq ID No: 444 & 445	447033	AI357412	Hs.157601	ESTs
	Seq ID No: 446 & 447	115522	BE614387	Hs.333893	c-Myc target JPO1
	Seq ID No: 448 & 449	410418	D31382	Hs.63325	transmembrane protease, serine 4
	Seq ID No: 450 & 451	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
85	Seq ID No: 452 & 453	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119
	Seq ID No: 454 & 455	452461	N78223	Hs.108106	transcription factor
	Seq ID No: 456 & 457	412420	AL035668	Hs.73853	bone morphogenetic protein 2
	Seq ID No: 458 & 459	416658	U03272	Hs.79432	fibrillin 2 (congenital contractural ara
	Seq ID No: 460 & 461	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antag

	Seq ID No: 462 & 463	437852	BE001836	Hs.256897	ESTs, Weakly similar to dJ365O12.1 [H.s.a
	Seq ID No: 464 & 465	402075			ENSP00000251056: Plasma membrane calcium
	Seq ID No: 466 & 467	421110	AJ250717	Hs.1355	cathepsin E
5	Seq ID No: 468 & 469	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 470 & 471	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 472 & 473	451668	Z43948	Hs.326444	cartilage acidic protein 1
	Seq ID No: 474 & 475	422282	AF019225	Hs.114309	apolipoprotein L
	Seq ID No: 476 & 477	425862	AK001504	Hs.159651	death receptor 6, TNF superfamily member
10	Seq ID No: 478 & 479	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),
	Seq ID No: 480 & 481	427747	AW411425	Hs.180655	serine/threonine kinase 12
	Seq ID No: 482 & 483	420281	AI623693	Hs.323494	Predicted cation efflux pump
	Seq ID No: 484 & 485	405932			C15000305:gi 3806122 gb AAC69198.1  (AF0
	Seq ID No: 486 & 487	405932			C15000305:gi 3806122 gb AAC69198.1  (AF0
15	Seq ID No: 488 & 489	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane
	Seq ID No: 490 & 491	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy
	Seq ID No: 492 & 493	417079	U65590	Hs.81134	interleukin 1 receptor antagonist
	Seq ID No: 494 & 495	430890	X54232	Hs.2699	glypican 1
	Seq ID No: 496 & 497	419721	NM_001650	Hs.288650	aquaporin 4
20	Seq ID No: 498 & 499	444471	AB020684	Hs.11217	KIAA0877 protein
	Seq ID No: 500 & 501	413063	AL035737	Hs.75184	chitinase 3-like 1 (cartilage glycoprote
	Seq ID No: 502 & 503	433800	AI034361	Hs.135150	lung type-I cell membrane-associated gly
	Seq ID No: 504 & 505	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
	Seq ID No: 506 & 507	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro
	Seq ID No: 508 & 509	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitte
25	Seq ID No: 510 & 511	410407	X65839	Hs.63287	carbonic anhydrase IX
	Seq ID No: 512 & 513	309931	AW341683		gbchd13d01.x1 Soares_NFL_T_GBC_S1 Homo s
	Seq ID No: 514 & 515	412719	AW016610	Hs.816	ESTs
	Seq ID No: 516 & 517	417034	NM_006183	Hs.80962	neurotensin
30	Seq ID No: 518 & 519	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam
	Seq ID No: 520 & 521	413753	U17760	Hs.75517	laminin, beta 3 (nicein (125kD), kalini
	Seq ID No: 522 & 523	425650	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen
	Seq ID No: 524 & 525	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage
	Seq ID No: 526 & 527	418663	AK001100	Hs.41690	desmocollin 3
35	Seq ID No: 528 & 529	418663	AK001100	Hs.41690	desmocollin 3
	Seq ID No: 530 & 531	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas
	Seq ID No: 532 & 533	406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad
	Seq ID No: 534 & 535	431846	BE019924	Hs.271580	uroplakin 1B
	Seq ID No: 536 & 537	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL
40	Seq ID No: 538 & 539	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta
	Seq ID No: 540 & 541	437044	AL035864	Hs.69517	differentially expressed in Fanconi's an
	Seq ID No: 542 & 543	428484	AF104032	Hs.184601	solute carrier family 7 (cationic amino
	Seq ID No: 544 & 545	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3
	Seq ID No: 546 & 547	417389	BE260964	Hs.82045	midkine (neurite growth-promoting factor
45	Seq ID No: 548 & 549	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3
	Seq ID No: 550 & 551	417542	J04129	Hs.82269	progestagen-associated endometrial prola
	Seq ID No: 552 & 553	449230	BE613348	Hs.211579	melanoma cell adhesion molecule
	Seq ID No: 554 & 555	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
	Seq ID No: 556 & 557	410555	U92649	Hs.64311	a disintegrin and metalloproteinase doma
50	Seq ID No: 558 & 559	424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B
	Seq ID No: 560 & 561	418462	BE001596	Hs.85266	integrin, beta 4
	Seq ID No: 562 & 563	410274	AA381807	Hs.61762	hypoxia-inducible protein 2
	Seq ID No: 564 & 565	439606	W79123	Hs.58561	G protein-coupled receptor 87
	Seq ID No: 566 & 567	404877			NM_005355: Homo sapiens melanoma antigen,
55	Seq ID No: 568 & 569	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote
	Seq ID No: 570 & 571	418543	NM_005329	Hs.85962	hyaluronan synthase 3
	Seq ID No: 572 & 573	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 574 & 575	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 576 & 577	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
60	Seq ID No: 578 & 579	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 580 & 581	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 582 & 583	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t
	Seq ID No: 584 & 585	421817	AF146074	Hs.108660	ATP-binding cassette, sub-family C (CFTR
	Seq ID No: 586 & 587	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
65	Seq ID No: 588 & 589	418678	NM_001327	Hs.167379	cancer/testis antigen (NY-ESO-1)
	Seq ID No: 590 & 591	409420	Z15008	Hs.54451	laminin, gamma 2 (nicein (100kD), kalini
	Seq ID No: 592 & 593	332180	AF134160	Hs.7327	claudin 1
	Seq ID No: 594 & 595	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
	Seq ID No: 596 & 597	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,
70	Seq ID No: 598 & 599	439223	AW238299	Hs.250618	UL16 binding protein 2
	Seq ID No: 600 & 601	409757	NM_001898	Hs.123114	cystatin SN
	Seq ID No: 602 & 603	428969	AF120274	Hs.194689	artemin
	Seq ID No: 604 & 605	428969	AF120274	Hs.194689	artemin
	Seq ID No: 606 & 607	428969	AF120274	Hs.194689	artemin
75	Seq ID No: 608 & 609	428969	AF120274	Hs.194689	artemin
	Seq ID No: 610 & 611	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 612 & 613	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-
	Seq ID No: 614 & 615	414774	X02419	Hs.77274	plasminogen activator, urokinase
	Seq ID No: 616 & 617	407944	R34008	Hs.239727	desmocollin 2
	Seq ID No: 618 & 619	407944	R34008	Hs.239727	desmocollin 2
80	Seq ID No: 620 & 621	457489	AI693815	Hs.127179	cryptic gene
	Seq ID No: 622 & 623	429547	AW009166	Hs.99376	ESTs
	Seq ID No: 624 & 625	407242	M18728		gbt:Human nonspecific crossreacting anti
	Seq ID No: 626 & 627	407242	M18728		gbt:Human nonspecific crossreacting anti
	Seq ID No: 628 & 629	407242	M18728		gbt:Human nonspecific crossreacting anti
85	Seq ID No: 630 & 631	444005	BE395085	Hs.10086	type I transmembrane protein Fn14

Seq ID No: 632 & 633	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase domain
Seq ID No: 634 & 635	422109	S73265	Hs.1473	gastrin-releasing peptide
Seq ID No: 636 & 637	419235	AW470411	Hs.288433	neurotrophin
Seq ID No: 638 & 639	449048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc
Seq ID No: 640 & 641	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy
Seq ID No: 642 & 643	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu
Seq ID No: 644 & 645	448243	AW369771	Hs.52620	integrin, beta 8
Seq ID No: 646 & 647	426427	M86699	Hs.169840	TTK protein kinase
Seq ID No: 648 & 649	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6
Seq ID No: 650 & 651	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 8
Seq ID No: 652 & 653	428450	NM_014791	Hs.184339	KIAA0175 gene product
Seq ID No: 654 & 655	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,
Seq ID No: 656 & 657	453392	U23752	Hs.32964	SRY (sex determining region Y)-box 11
Seq ID No: 658 & 659	426514	BE616533	Hs.170195	bone morphogenetic protein 7 (osteogenic
Seq ID No: 660 & 661	425776	U25128	Hs.159499	parathyroid hormone receptor 2
Seq ID No: 662 & 663	425776	U25128	Hs.159499	parathyroid hormone receptor 2
Seq ID No: 664 & 665	431515	NM_012152	Hs.258583	endothelial differentiation, lysophospha
Seq ID No: 666 & 667	419452	U33635	Hs.90572	PTK7 protein tyrosine kinase 7
Seq ID No: 668 & 669	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
Seq ID No: 670 & 671	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
Seq ID No: 672 & 673	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
Seq ID No: 674 & 675	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci
Seq ID No: 676 & 677	410001	AB041036	Hs.57771	kallikrein 11
Seq ID No: 678 & 679	426501	AW043782	Hs.293616	ESTs
Seq ID No: 680 & 681	408369	R38438	Hs.182575	solute carrier family 15 (H777) transport
Seq ID No: 682 & 683	445413	AA151342	Hs.12577	CGI-147 protein
Seq ID No: 684 & 685	422424	AJ186431	Hs.295638	prostate differentiation factor
Seq ID No: 686 & 687	428330	L22524	Hs.22556	matrix metalloproteinase 7 (matrilysin,
Seq ID No: 688 & 689	420610	AJ683183	Hs.99348	distal-less homeo box 5

TABLE 15B

Pkey: Unique Eos probe/identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT Number	Accession
309931	AW341683	
330493	33264_5	M27826 R78416 AA307645 AW957879 AW957800 AA633529 H03682
439285	47065_1	AL133916 N79113 AF086101 N76721 AW950828 AA364013 AW955684 AJ346341 AJ867454 N54784 AJ655270 AJ421279 AW014882
		AA775552 N62351 N59253 AA626243 AJ341407 BE175639 AA456968 AJ358918 AA457077
450375	83327_1	AA009647 AA131254 AA374293 AW954405 H04410 AW606284 AA151166 BE157457 BE157601 H04384 W46291 AW663674 H04021 H01532
		AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AJ161014 AA099554 R69067
451320	86576_1	AW118072 AJ631982 T15734 AA224195 AJ701458 W20198 F26326 AA890570 N90552 AW071907 AJ671352 AJ375892 T03517 R88265
		AJ124088 AA224388 AJ084316 AJ354686 T33652 AJ140719 AJ720211 T03490 AJ372637 T15415 AW205836 AA630384 T03515 T33230
		AA017131 AA443303 T33623 AJ222556 T33511 T33785 AJ419606 D55612

TABLE 15C

Pkey: Unique number corresponding to an Eos probe/identifier  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
403329	8516120	Plus	96450-96598
403478	9958258	Plus	116458-116564
404440	7528051	Plus	80430-81581
404877	1519284	Plus	1095-2107
405770	2735037	Plus	61057-62075
405932	7757812	Minus	123525-123713

Table 16

Seq ID NO: 1 DNA sequence  
Nucleic Acid Accession #: NM\_001216  
Coding sequence: 43..1422

	1	11	21	31	41	51	
5							
10	GCCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCATGGCTCC	CCTGTGCCCC	60
	AGCCCCCTGGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGCG	TGCTTCTGAT	GCCTGTCCAT	CCCCAGAGGT	TGCCCGGAT	GCAGGAGGAT	180
	TCCCCCTTGG	GAGGAGGCTC	TTCTGGGAA	GATGACCCAC	TGGCGAGGA	GGATCTGCCC	240
	AGTGAAGAGG	ATTCAACCAG	AGAGGAGGAT	CCACCCGGAG	AGGAGGATCT	ACCTGGAGAG	300
15	GAGGATCTAC	TGGGAGAGGA	GGATCTACCT	GAAGTTAAGC	CTAATCAGA	AGAAGAGGGC	360
	TCCCTGAAGT	TAGAGGATCT	ACCTACTGTT	GAGGCTCCTG	GAGATCTCTA	AGAACCCAG	420
	AATAATGCCC	ACAGGGACAA	AGAAGGGGAT	GACCAGAGTC	ATTGGCGCTA	TGGAGGCGAC	480
	CGCCCTTGGC	CCCGGGTGTG	CCCAGCCTGC	GCGGGCCGCT	TCCAGTCCCC	GGTGGATATC	540
	CGCCCCAGC	TCGCGCGCTT	CTGCCCGGCC	CTGCCGCCCC	TGGAATCCTT	GGGCTTCCAG	600
20	CTCCCGCCGC	TCCCAGAACT	GCGCTTGGC	AACAATGGCC	ACAGTGTGCA	ACTGACCCCTG	660
	CCTCTCTGGC	TAGAGATGGC	TCTGGGTCCC	GGCGGGAGT	ACCGGCTCT	GCAGCTGCAT	720
	CTGCACTGGG	GGCTGTGAGG	TGTTCCGGGC	TGAGAGCACA	CTGTGGAAGG	CCACCGTTTC	780
	CCTGCCGAGA	TCCAGTGGT	TCACTCAGC	ACCGCTTTG	CCAGAGTTGA	CGAGGCTTG	840
	GGGCGCCCGG	GAGGCTTGGC	CGTGTGGCC	GCCTTTCTGG	AGGAGGCCCC	GGAGAGAAAC	900
25	AGTGCCTATG	AGCAGTTGCT	GTCTCGCTTG	GAAGAAATCG	CTGAGGAAGG	CTCAGAGACT	960
	CAGGTCCAG	GACTGGACAT	ATCTGCATCT	CTGCCCTCTG	ACTTCAGCCG	CTACTTCCAA	1020
	TATGAGGGGT	CTCTGACTAT	ACCGCCTGT	GCCAGGGTG	TCATCTGGAC	TGTGTTTAAC	1080
	CAGACAGTGA	TGCTGAGTGC	TAAGCAGCTC	CACACCTCT	CTGACACCT	GTGGGACCT	1140
	GGTGACTCTC	GGCTACAGCT	GAACCTCCGA	GCGACGACG	CTTTGAATGG	GCGAGTGATT	1200
30	GAGGCTCTCT	TCCCTGCTGG	AGTGGACAGC	AGTCTCTCGG	CTGCTGAGCC	AGTCCAGCTG	1260
	AATTCTCTCC	TGGCTGTGG	TGACATCTTA	GCCCTGGTTT	TTGGCTCTCT	TTTGTCTGTC	1320
	ACCAGCGTCG	CGTCTCTTGT	GCAGATGAGA	AGGCAGCACA	GAAGGGGAAC	CAAAAGGGGT	1380
	GTGAGCTACC	GCCACAGAGA	GGTAGCCGAG	ACTGGAGCCT	AGAGGCTGGA	TCTTGGAGAA	1440
	TGTGAGAAGC	CAGCCAGAGG	CATCTGAGGG	GGAGCCGCTA	ACTGTCTGT	CCTGCTCATT	1500
35	ATGCCACTTC	CTTTAACTG	CCAAGAAAT	TTTTAAATA	AATATTATA	AT	

Seq ID NO: 2 Protein sequence:  
Protein Accession #: NP\_001207

	1	11	21	31	41	51	
40							
	MAPLCPSFWL	PLLIPAPAPG	LTVQLLSLL	LMPVHPQRL	PRMQEDSPLG	GGSSGEDDPL	60
	GEEDLPSSED	SPREDPPGE	EDLPGEEDLP	GEEDLPVVKP	KSEEGSLKL	EDLPTVEAPG	120
	DPQEPQNAH	RDKGGDDQSH	WRYGGDFPWP	RVPACAGRF	QSFVDIRPOL	AAFCPALRPL	180
45							
	ELLGFLPPL	PELRLENNGH	SVQLTLPPL	EMALGPQREY	RALQLHLHWG	AAGRPGSEET	240
	VEGHRFPFBI	HVHLSTAF	RVDEALRPG	GLAVLAAPLE	EGPEENSAYE	QLLSRLIEIA	300
	EGSSETQVFG	LDSALLPSG	FSRYFQYEGS	LTPPCAQGV	IWTVFNQTM	LSAKQLHTLS	360
	DTLWGPDSR	LQLNFRATQP	LNGRVIEASF	PAGVDSPPRA	AEPVQLNSCL	AAGDILALVF	420
50							
	GLLPAVTSVA	FLVQMRQRH	RGTGGVSYSR	PAEVAETGA			

Seq ID NO: 3 DNA sequence  
Nucleic Acid Accession #: BC013923  
Coding sequence: 438-1391

	1	11	21	31	41	51	
55							
	AGCGGGGTG	TCTATTAAC	TGTTCAAAA	GTATCAGGAG	TGTCAAGGC	AGAGAAGAGA	60
	GTGTTTGCAA	AAGGGGAA	GTAGTTTGT	GCCTCTTAA	GACTAGGACT	GAGAGAAAGA	120
	AGAGGAGAGA	GAAAGAAAGG	GAGAGAAGTT	TGAGCCCCAG	GCTTAAGCCT	TTCCAAAAA	180
60							
	TAATAATAAC	AATCATCGC	GGCGGCAGGA	TGGCCAGAG	GAGGAGGGAA	GGCTTTTTT	240
	TGATCTGAT	TCCAGTTTGC	CTCTCTCTT	TTTTCCCCCA	AATTATTCTT	CGCTGATT	300
	TCTCTCGGGA	GCCTTGCCT	CCGACACCC	CGCCCGCCT	CCCTCTCTC	TCTCCCCCG	360
	CCCGCGGGC	CCCCAAAGT	CCGCGCGGC	CGAGGTCGG	CGGCGCGCG	CGGCGCGGC	420
	CCGCGCACAG	CGCCCGCATG	TACAACATGA	TGGAGACGGA	GCTGAAGCG	CCGCGCGGC	480
65							
	AGCAAACTTC	GGGGGCGGC	GGCGCAACT	CCACGCGGC	GGCGCGCGC	GGCAACAGGA	540
	AAAACAGCCC	GGACCGCTC	AAGCGGCCA	TGAATGCCTT	CATGGTGTGG	TCCCGCGGC	600
	AGCGGCGCAA	GATGGCCAG	GAGAACCCCA	AGATGCACAA	CTCGAGATC	AGCAAGCGCC	660
	TGGGCGCGGA	GTGAAACTT	TGTTCGAGA	CGGAGAAGCG	CGCGTTCATC	GACGAGGCTA	720
70							
	AGCGGCTGCG	AGCGCTGCAC	ATGAAGGAGC	ACCGGATTA	TAAATACCG	CCCGGCGGA	780
	AAACCAAGAC	GCTCATGAAG	AAGGATAAGT	ACACGCTGCC	CGCGGGCTG	CTGGCCCCCG	840
	CGCGCAATAG	CATGCGGAGC	GGGTGCGGG	TGGGCGCGG	CCTGGGCGCG	GGCGTGAACC	900
	AGCGCATGGA	CAGTTACGCG	CACATGAACG	GCTGGAGCAA	CGGCAGCTAC	AGCATGATGC	960
	AGGACAGCT	GGGCTACCG	CAGCACCCGG	GCCTCAATGC	GCACGGCGCA	GCGCAGATGC	1020
	AGCCCATGCA	CGGCTACGAC	GTGAGCGCC	TGCAGTACAA	CTCCATGACC	AGCTCGCAGA	1080
75							
	CCTACATGAA	CGGCTCGCC	ACCTACAGCA	TGCTCTACTC	GCAGCAGGCG	ACCCCTTGCA	1140
	TGGCTCTTGG	CTCCATGGGT	TGGTGGTCA	AGTCCGAGC	CAGCTCCAGC	CCCCCTGTGG	1200
	TTACCTCTTC	CTCCACTCC	AGGGCGCCCT	GCCAGGCGCG	GGACCTCCGG	GACATGATCA	1260
	GCATGTATCT	CCCCGCGCC	GAGGTGCCGG	AACCGCGCG	CCCCAGCAGA	CTTCACATGT	1320
	CCAGCACTA	CCAGAGCGGC	CCGTGCCCG	GCACGGCAT	TACGCGACA	CTGCCCTCT	1380
	CACACATG	AGGGCGGAC	AGCGAACTGG	AGGGGGGAGA	AATTTTCAA	GAAAAACGAG	1440
80							
	GGAAATGGGA	GGGTGCAAA	AGAGGAGAGT	AAGAAACAGC	ATGGAGAAAA	CCCGGTACGC	1500
	TCAAAAAAAA	AAAAAATAA	AAATCCCAT	CACCCACAGC	AAATGACAGC	TGCAAAAGAG	1560
	AACCACAATC	CCATCCACAC	TCAAGCAAAA	ACCGCGATGC	CGACAAGAAA	ACTTTTATGA	1620
	GAGAGATCTT	GGACTTCTTT	TKGGGGGACT	ATTTTGTAC	AGAGAAAAAC	TGGGGAGGGT	1680
	GGGAGGGCG	GGGGAATGGA	CCTTGTATAG	ATCTGGAGGA	AAGAAAGCTA	CGAAAACTT	1740
85							
	TTTAAAGTT	CTAGTGGTAC	GGTAGAGCT	TTGCAGGAAG	TTTGCAAAAG	TCTTTACCAA	1800
	TAATATTAG	AGCTAGTCTC	CAAGCGACGA	AAAAAATGTT	TTAATATTG	CAAGCAACTT	1860
	TGTACAGTA	TTTATCGAGA	TAAACATGCG	AATCAAAATG	TCCATTGTTT	ATAAGCTGAG	1920

AATTGGCCAA TATTTTTCAA GGAGAGGCTT CTTGCTGAAT TTTGATCTG CAGCTGAAAT 1980  
 TTAGGACAGT TGCAAAAGTG AAAAGAGAGAA AATTATTCAA ATTTGGACAT TTTAATTGTT 2040  
 TAAAAATTGT ACAAAGGAA AAAATTAGAA TAAGTACTGG CGAACCATCT CTGTGGTCTT 2100  
 GTTTAAAAAG GGCAAAAGTT TTAGACTGTA CTAATTTTA TAACCTACTG TTAAGAGCAA 2160  
 AAATGGCCAT GCAGGTTGAC ACCGTGGTA ATTTATAATA GCTTTGTTC GATCCCACT 2220  
 TTCCATTTTG TTCAGATAAA AAAAACCATG AAATTACTGT GTTTGAAATA TTTCTTATG 2280  
 GTTTGTAATA TTTCTGTAAA TTTATTGTGA TATTTTAAGG TTTTCCCCC TTTATTTTCC 2340  
 GTAGTTGTAT TTTAAAGAT TCGGCTCTGT ATTATTGAA TCAGTCTGCC GAGAAATCCAT 2400  
 GTATATATT GAACATAAT CATCCTTATA ACAGGTACAT TTTCAACTTA AGTTTITACT 2460  
 CCATTATGCA CAGTTTGAGA TAAATAAAT TTTGAAATAT GGACACTGAA AAAAAAAGAA 2520  
 AAAAAACAA ACAAAGAAA CAAAAACAA AAACAGAAAA AACAAAAAA AACAACAAAC 2580  
 CACAACACAA AAACAACAAA AAAAAAAGA AACAAACACA CAACACACA CAACACAAAA 2640  
 CCACAACACA AACACACAA CACAGAGGG

Seq ID NO: 4 Protein sequence:  
Protein Accession #: CAA83435.1

1 11 21 31 41 51  
 MYNMMETELK PPGPQQTSGG GGGNSTAAAA GGNQKNSPDR VKRPMNAFV WSRGQRRKMA 60  
 QENPKMENSE ISKRLGAENK LLSETEKRPF IDEAKRLRAL HMKHEPDYKY RPRRKTKTLM 120  
 KKDKYTLPGG LLAPGGSMA SGVGVGAGLG AGVNQRMDSY AHMNGWSNGS YSMMDOLGY 180  
 PQHPGLNAGS AAQMPMERY DVSLQYNSM TSSQTYMNGS PTYSMSYSQQ GTPGMALGSM 240  
 GSVVKSEASS SPPVVTSSSH SRAPCQAGDL RDMISMVLPF AEVPEPAAPS RLHMSQHYQS 300  
 GPVPGTAING TLPLSEH

Seq ID NO: 5 DNA sequence  
Nucleic Acid Accession #: U91618  
Coding sequence: 29-541

1 11 21 31 41 51  
 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60  
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAAATGAA 120  
 AGCATTAGAA GCAGATTCTT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180  
 TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCACTCTT GTAAATAATT TGAACAGCCC 240  
 AGCTGAGGAA ACAGGAGAAG TTCTATGAAG GGAGCTTGTT GCAAGAAGGA AACTTCTTAC 300  
 TGCTTTAGAT GGCTTTAGCT TGGAAAGCAAT GTTGACAATA TACCAGCTCC ACAAATCTG 360  
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAA 420  
 TGACAAAAT GGAAAGGAAG AAGTCATAAA GAGAAAAAT CCTTATATTC TGAACGCGCA 480  
 GCTGTATGAG AATAAACCCA GAAGACCTTA CATACTCAA AGAGATTCTT ACTATTACTG 540  
 AGAGAAATAA TCATTATTT ACATGTGATT GTGATTCAT ATCCCTTAAT TAAATATCAA 600  
 ATTATATTG TGTGAAAATG TGACAAACAC ACTTATCTGT CTCCTTACA ATTGTGGTTT 660  
 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720  
 TCTTCAAAAA AAAAAAATAA AAATGGGGCC GCAATT

Seq ID NO: 6 Protein sequence:  
Protein Accession #: AAB50564

1 11 21 31 41 51  
 MMAGMKIQLV CMLLAPSSW SLCSDSSEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60  
 VCSLVNNLNS PAETGEVHE BELVARKLP TALDGFSLBA MLTIYQLHKI CHSRAPQHWZ 120  
 LIQEDILDGT NDKNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSYFY

Seq ID NO: 7 DNA sequence  
Nucleic Acid Accession #: NM\_006536.2  
Coding sequence: 109-2940

1 11 21 31 41 51  
 ACCTAAAAAC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60  
 ATGTATGCRG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACACAT GACCCAAAGG 120  
 AGCATTGCAG GTCTATTGTC CAACCTGAAG TTTGTGACTC TCCTGGTTGC CTTAAGTTCA 180  
 GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATTA TGGATTGCTC 240  
 ATTGCAATTA ATCTCTCAGT ACCTGAGAAT CAGAACCTCA TCTCAACAT TAAGGAAATG 300  
 ATAACTGAAG CTTCAATTTA CCTATTAAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360  
 ATAAAGATT TAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAGAAA 420  
 TCATATGAAA AGGCAAAATG CATAGTGACT GACTGGTATG GGGCACATGG AGATGATCCA 480  
 TACACCTTAC AATACAGAGG GTGTGGAAAA GAGGGAAAAAT ACATTCATTT CACACCTAAT 540  
 TTCTACTGTA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCGGAGT GTTTGTCCAT 600  
 GAATGGGCCC ACCTCCGTTG GGGTGTGTTT GATGAGTATA ACAATGACAA ACCTTCTTAC 660  
 ATAAATGGGC AAAATCRAAT TAAAGTGACA AGGTGTTTCA CTGACATCAC AGGCATTTTT 720  
 GTGTGTGAAA AAGTCTCTTG CCCCAAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780  
 GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840  
 AGTTTATCTT CTGTGGTTGA ATTTTGTAAAT GCAAGTACCC ACAACCAAGA AGCACCACAA 900  
 CTACAGAAC AGATGTGAGC CCTCAGAAAT GCATGGGATG TAATCACAGA CTCTGCTGAC 960  
 TTTCAACACA GCTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCCAC ATTCTCGCTT 1020  
 GTACAGGCTG GTGACAAAGT GGTCTGTTTA GTGCTGGATG TGTCCAGCAA GATGCGCAG 1080  
 GCTGACAGAG TCCTTCAACT ACAACAAGCC GCAGAAATTT ATTTGATGCA GATGTGTGAA 1140  
 ATTCAATCT TCGTGGGCAAT TGCCAGTTTC GACAGCAAGG GAGAGATCAG AGCCCGACTA 1200  
 CACCAATTA ACAGCAATGA TGATCGAAGG TTGCTGTTT CATATCTGCC CACCACCTGA 1260  
 TCAGCTAAAA CAGATATCA CATTGTGTTA GGGCTTAAGA AAGGATTTGA GGTGGTTGAA 1320  
 AAACCTGAAT GAAAGCTTA TGGCTCTGTG ATGATATTAG TGACCAGCGG AGATGATAAG 1380  
 CTTCTTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCTATTGCC 1440

CTGGGTTTCAT CTGCAGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500  
 TTCTTTGTTC CAGATATATC AAATCCCAAT AGCATGATTG ATGCTTTTCAG TAGAATTTCC 1560  
 TCTGGAACTG GAGACATTTT CCAGCAACAT ATTCACTCTG AAAAGTACAGG TGAAAAATGTC 1620  
 AAACCTCAACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680  
 5 ATGTTTCTAG TTACGTGGCA GGCACAGTGGT CCTCTGAGA TTATATTATT TGATCCTGAT 1740  
 GGAACGAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800  
 TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTGGAACAA TACCCATCAT 1860  
 TCTCTGCAAG CCTCGAAAGT GACAGTGACC TCTCGCGCCT CCAACTCAGC TGTGCCCCCA 1920  
 10 GOCACGTGGG AAGCCTTTGT GGAAAGAGAC AGCCTCCATT TTCCTCATCC TGTGATGATT 1980  
 TATGCCAATG TGAACACGGG ATTTTATCCC ATCTTAAATG CCACTGTCCAC TGCCACAGTT 2040  
 GAGCCAGAGA CTGGAGATCC TGTTAGCCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100  
 GTTATAAAAA ATGATGGAAT TTAATCGAGG TATTTTCTCT CCTTTCCTGC AAATGGTAGA 2160  
 TATAGCTTGA AAGTGCAATG CAATCACTCT CCCAGCATAA GCACCCCAAC CCACTCTATT 2220  
 CCAGGGAGTC TAGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280  
 15 GCTCCAGGA ATACAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340  
 AGCTCAGGAG GCTCCTTTTC AGTCTGGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400  
 CCACCATGCA AAATTATTGA CTGGAAGCT GTAAAGTAG AAGAGGAATT GACCCATCT 2460  
 TGGACAGCAC CTGGAGAAGA CTTTGATCAG GGCCAGGCTA CAAGCTATGA AATAAGAATG 2520  
 AGTAAAGTC TACGAATATG CCAAGATGAC TTTAACATG CTATTTTAGT AAATACATCA 2580  
 20 AAGCGAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA CGTCTCACC CCAGATTTC 2640  
 ACGAATGAGC CTGAACATCA GCCAAATGGA GAAACACATG AAGCCACAG AATTATGTT 2700  
 GCAATACGAG CAATGGATAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCACGCG 2760  
 CCTCTGTTTA TCCCCCCTTA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820  
 GGAATTTTAA GAGCAATGGG TTTGATAGGA ATCATTGCCC TTATTTAGT TGTGACAT 2880  
 25 CATACTTTAA CGAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATA 2940  
 ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAA 3000  
 CATACTAACA AAGTCAATT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060  
 ATACAGATAA GATTTTATCA TGGTAGATCA ACAATCTTCT TTGGGGGTAG ATTAGAAAA 3120  
 CCTTACACTT TGGCTATGAA CAAATAATAA AAATTATTCT TTAAGTAAT GTCTTTAAAG 3180  
 30 GCAAGGGGAA GGGTAAAGTC GGACCAAGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAA 3240  
 AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACGTCTG TGTGAAGCAA 3300  
 TCATTITAGT ACTTTGATTA ATTTTCTTCT TCTCCTTATC TGTGCAGTAC AGGTTGCTTG 3360  
 TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAG CTCTTTACCT 3420  
 35 CTTCCTATT TGTATATAT ATTTACAGAT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480  
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540  
 TTTATGACAA AGGCTCATTT AATTTATTG TGTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600  
 TTTCTAGATT TATTGCTTGT GGTATTATG GAATGATAGT TATAGCCCN TATAATGCTT 3660  
 TACCTAGGAA A

Seq ID NO: 8 Protein sequence:  
 Protein Accession #: NP\_006527.1

1 11 21 31 41 51  
 45 | | | | |  
 MTQRSIAGPI CNLKFVTLV ALSSELPLFG AGVQLQDNGY NGLLIAINPQ VPENQNLISS 60  
 IKEMITEASF YLFNATKRRV FFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAH 120  
 GDDPYTLQYR GCGKGGKYIH FTFNLLNDN LTAGYGSRRR VVVEWAHLR WGVFDEYNND 180  
 KPFYINGNQI IKVTSCSDI TGI FVCEKGP CPOENCIISK LFKEGCTFIY NSTQNTASI 240  
 50 MPMQSLSSVV EFCNASTENQ EAPNLQNMOC SLRSANDVIT DSADFHHSFP MNGTELPPPP 300  
 TFSLVQAGDK VVCLVLDVSS KMAEADRLQ LQQAEEFLYM QIVEIHTFVG IASFDSKGEI 360  
 RAQLHQINSN DDKLLVSVL PTTVSAKTDI SICSLKKGF EVVEKLNGKA YGSMVLVTS 420  
 GDDKLLQNLCL FTVLSSGSTI HSIALGSSAA PNLEELSRIT GGLKFFVPDI SNSNSMIDAF 480  
 SRISSTGDI DFDQGLST GENVKPEQL KNTVTVDNTV GNDTFLVTV QASGPPEILL 540  
 55 FDPDGRKYIT NNFITNLTR TASLWIPGTA KPGHWYITLN NTHBSIQALK VVTSTRASNS 600  
 AVPPATVEAF VERDSLHFP PVMYANVKQ GFYPIILNATV TATVEPETGD PVTLRLLDDG 660  
 AGADVIKNDG IYSRYFFSFA ANGRYSKLVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720  
 IQMNAFRKSV GRNEEERKVG FSRVSSGGSF SVLGVPAGPH PDVFPCKII DLEAVKVEEE 780  
 LTLSTWAPGE DFDQGAQTSY EIRMSKSLON IQDDFNAIL VNTSKRNPOQ AGIREIFTFS 840  
 60 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900  
 LILKGVLTAM GLIGIICLII VVTHHTLSRK KRADKKENG T KLL

Seq ID NO: 9 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 336-632

1 11 21 31 41 51  
 70 | | | | |  
 CTCCCTCAC CCGGGTCCAG GATGCCAGT CCCCACGACA CCTCCACTT CCCACTGTGG 60  
 CTGGGGTGGG CTCAGGGGCT GCCCTTGACC TGGCCTAGAG CCTCCCCCA GCTGGTGGTG 120  
 GAGCTGGCAC TCTCTGGGAG GGAGGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180  
 CAGGGTTTGG TGGATCAGG TTGAGGCAGG TTTGTTTTC TTAATGTC AAGTTGGGGG 240  
 CCAGTGGGGC CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCTTTC TCTCCTTCT 300  
 75 GGGTCTGTCT CTGCCACCTG GTCTGCCACA GATCCATGAT GTGCAGTCT CTGAGCAGG 360  
 CGCTGGCTGT GCTGCTCACT ACCTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420  
 AGCTGAGTAA GGGGGAART AAGGAACCTC TGCACAAGGA GCTGCCAGC TTTGTGGGG 480  
 AGAAAGTGA TGAAGAGGGG CTGAAGAAGC TGATGGGAG CCTGGATGAG AACAGTGACC 540  
 AGCAGGTGA AGCTCAGGAG TATGCTGTTT TCCTGGCACT CATCACTGTC ATGTGCAATG 600  
 80 ACTTCTTCCA GGGCTGCCCA GACCGACCTT GAAGCAGAAC TCTTGACTTC CTGCCATGGA 660  
 TCTCTGGGC CCAGGACTGT TGATGCCCTT GAGTTTGTGA TTCAATAAAC TTTTGTGTC 720  
 TGTGATTAAT ATTTTAATG CTGAGTGATG TTCCATAACC CGGCTGGCTC AGCTGAGTG 780  
 CTGGGAGATG AGGGCCTCCT GGATCCTGCT CCTTCTGGG CTCTGACTCT CCTGGAATC 840  
 TCTCCAGGC GAGGACTATG CTTTAGGCTT CAATTTTGA ATTTCAACA CCAGCAAAA 900  
 85 ATTGGAATC GAGATAGGTT GCTGACTTTT ATTTTGTCAA ATAAAGATAT TAAAAAGGC 960  
 AATACCA

Seq ID NO: 10 Protein sequence:

Protein Accession #: NP\_005969.1

1 11 21 31 41 51  
5 | | | | |  
MMCSLEQAL AVLVTTPHKY SQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60  
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

Seq ID NO: 11 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 336-626

1 11 21 31 41 51  
15 | | | | |  
CTCCCTCAC CCGGTCAG GATGCCAGT CCCCACGACA CCTCCACTT CCACTGTGG 60  
CCTGGGTGGG CTGAGGGGCT GCCCTTGACC TGGCTAGAG CCTCCCCCA GCTGGTGGTG 120  
GAGCTGGCAC TCTCTGGGAG GGAGGGGCT GGGAGGGAAT GAGTGGGAAT GGCAAGAGGC 180  
CAGGGTTTGG TGGGATCAGG TTGAGGCAGG TTTGGTTTCC TTAAATGCC AAGTTGGGG 240  
20 CCACTGGGGC CCACATATAA ATCCTCACCC TGGGAGCCTG GCTGCCCTGC TCTCCTTCT 300  
GGGTCTGTCT CTGCCACTG GTCTGCCACA GATCCATGAT GTGCAGTTCT CTGGAGCAGG 360  
CGCTGGCTGT CTGGTCACT ACCTTCCACA AGTACTCCTG CCAAGAGGGC GACAAGTTCA 420  
AGCTGAGTAA GGGGAAATG AAGGAACCTC TGCACAAGGA GCTGCCAGC TTTGTGGGG 480  
ATTCCAGAGA ACCATGTGCT GTGAGGGCCT TCCGAGTCCA TCTGTTAAT CCTGTCATTG 540  
25 GAGACTTGAG AAACAGAGC CCAGAAGGGA AAAGTGATTG TCCCAAGATC ACACAGCACT 600  
GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC AGCCTGGATG AGAACAGTGA 660  
CAGCAGGTG GCTTCCAGT AGTATGCTGT TTTCTGGCA CTCATCACTG TCATGTGCAA 720  
TGACTTCTTC CAGGGCTGCC CAGACGACC CTGAAGCAGA ACTCTTGACT TCCTGCCATG 780  
GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG TATTCAATAA ACTTTTGTG 840  
30 TCTGTTGATA ATATTTTAAT TGCTCAGTGA TGTTCATAA CCGGCTGGC TCAGCTGGAG 900  
TGCTGGGAGA TGAGGGGCTC CTGGATCCTG CTCCTTCTG GGCTCTGACT CTCCTGAAA 960  
TCTCTCAAG GCCAGAGCTA TGCTTAGGT CTCATTTTG GAATTTCAA CACCAGCAA 1020  
AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGT AAATAAAGAT ATTAATAAAG 1080  
GCAATACCA

Seq ID NO: 12 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
40 | | | | |  
MMCSLEQAL AVLVTTPHKY SQEGDKFKL SKGEMKELLH KELPSFVGHS REPCAIVRAFR 60  
VHLFNPVIGD LRNQSPCKS DCPKITQHRW KWMRRG

Seq ID NO: 13 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 58-354

1 11 21 31 41 51  
50 | | | | |  
GTGAGCTCAC CATGTGGGGG TGAGGCTGAG AGAAAACAAG TACACAGCCA CAGATCCATG 60  
ATGTGCAATT CTCTGGAGCA GCGCTGGGCT GTGCTGGTCA CTACCTTCCA CAGTACTCC 120  
TGCCAGAGAG GCGACAAGTT CAAGCTGAGT AAGGGGGAAT TGAAGGAAGT TCTGCACAAG 180  
GAGCTGCCCA GCTTTGTGGG GGAGAAAGTG GATGAGGAGG GGCTGAAGAA GCTGATGGGC 240  
55 AGCCTGGATG AGAACAGTGA CCAGCAGGTG GACTTCCAGG AGTATGCTGT TTTCTGGCA 300  
CTCATCACTG TCATGTGCAA TGACTTCTTC CAGGGCTGCC CAGACGACC CTGAAGCAGA 360  
ACTCTTGACT TCCTGCCATG GATCTCTTGG GCCCAGGACT GTTGATGCCT TTGAGTTTGG 420  
TATTCAATAA ACTTTTGTG TCTGTTGATA ATATTTTAAT TGCTCAGTGA TGTTCATAA 480  
CCCGGCTGGC TCAGCTGAGG TGCTGGGAGA TGAGGGGCTC CTGGATCCTG CTCCTTCTG 540  
60 GGCTCTGACT CTCCTGAAA TCTCTCAAG GCCAGAGCTA TGCTTAGGT CTCATTTTG 600  
GAATTTCAA CACCAGCAA AAATTGAAA TCGAGATAGG TTGCTGACTT TTATTTTGT 660  
AAATAAAGAT ATTAATAAAG GCAATACCA

Seq ID NO: 14 Protein sequence:  
Protein Accession #: NP\_005969.1

1 11 21 31 41 51  
65 | | | | |  
MMCSLEQAL AVLVTTPHKY SQEGDKFKL SKGEMKELLH KELPSFVGEK VDEEGLKKLM 60  
GSLDENSQQ VDFQYAVFL ALITVMCNDF FQGCPRDP

Seq ID NO: 15 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 62-358

1 11 21 31 41 51  
75 | | | | |  
GGAGGGTGTG CCGCTGAGTC ACTGCGTGGG CATCTGGGCC TGGAACTCG GCCACAGATC 60  
CATGATGTGC AGTTCTCTGG AGCAGGGGCT GGCTGTGCTG GTCACCTACT TCCACAAGTA 120  
CTCCTGCCAA GAGGGCGACA AGTTCAAGCT GAGTAAGGGG GAATTAAGAG AACTTCTGCA 180  
80 CAGAGAGCTG CCGAGCTTTG TGGGGGAGAA AGTGGATGAG GAGGGGCTGA AGAAGCTGAT 240  
GGCGAGCCTG GATGAGAACA GTGACCAGCA GGTGGACTTC CAGGAGTATG CTGTTTCTCT 300  
GGCACTCATC ACTGTCTAGT GCAATGACTT CTTCCAGGGC TGCCAGACC GACCCTGAAG 360  
CAGAACTCTT GACTTCTGTC CATGGATCTC TTGGGCCGAC GACTGTTGAT GCCCTTGAGT 420  
75 TTTGATTCA ATAACTTTT TTTGTCTGTT GATAATATTT TAATGTCTCA GTGATGTTCC 480  
ATAAGCCGGC TGGCTCAGCT GGAGTGCTGG GAGATGAGGG CCTCCTGGAT CCTGCTCCT 540  
85 TCTGGGCTCT GACTCTCCTG GAAATCTCTC CAAGGCCAGA GCTATGCTTT AGGTCTCAAT 600  
TTTGAATTT CAAACACCAG CAAAAATTG GAAATCGAGA TAGGTTGCTG ACTTTTATTT 660

WO 02/086443

TGTCAAATAA AGATATTAAA AAAGGCAAAT ACCA

PCT/US02/12476

Seq ID NO: 16 Protein sequence:

Protein Accession #: NP\_005969.1

1 11 21 31 41 51  
| | | | |  
MMCSLSLEQAL AVLVTTPHKY SCQEGDKFKL SKGEMKELLE KELPSFVGEK VDEEGLKKLM 60  
GSLDENSDQQ VDFQYAVFL ALITVMCNDP FQGCPRDP

Seq ID NO: 17 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 939-2372

1 11 21 31 41 51  
| | | | |  
AAGACGGATT CTCAGACAAG GCTTGCAAAT GCCCGCAGC CATCATTTAA CTGCACCCGC 60  
AGAATAGTTA CGGTTTGTCA CCCGACCCCT CCGGATCGCC TAATTGTGTC CTAGTGAGAC 120  
CCCGAGGCTC TCCCGCGGCC TGGCTTCTTC GTAGCTGGAT GCATATCGTG CTCCGGGCAG 180  
CGCGGGCGCA GGGCAGCGGT TCGCGCACAC CCTAGCACAC ATGAACACGC GCAAGAGCTG 240  
AACCAAGCAC GGTTCCTATT TCAAAAAGGG AGACAGCCTC TACCGGATT GTAGAAGAGA 300  
CTGTGGTGTG AATTAGGAC CCGGAGGCGT CGAACGGAGG AACGGTTCAT CTTAGAGACT 360  
AATTTTCTGG AGTTTCTGCC CTGTCTCTGC GTACGCCCTC AGTCACTCTC GCCAGCAGTA 420  
GCAGAGGCGG CGCGGGCGGC TCCCGGAATT GGGTTGGAGC AGGAGCCTCG CTGGCTGCTT 480  
CGCTCGGCTC CTACGCGCTC AGTCCCGGCG GGTAGCAGGA GCCTGGACCC AGGCGCGGCC 540  
GGCGGGGCTG AGCGCGCGGA GCCCGGCCCTC GAGGTGCATA CGGACCCCTC ATTGCATCT 600  
AACAGGAAT CTGCCGCCCA GAGAGTCCCG GGAGCGCCGC CGGTCTGTGC CCGCGCGGCC 660  
GGGCCATGCA GGGACGGCCG CCGCGGAGCT CCGAGCAGCG GTAGCGCCCT CCTGTAAAGC 720  
GGTTCTGTAT GCGCGGGCCA CTGTGAACCC TGCCGCTGCG CGGAACACTC TTGCTCCGG 780  
ACCACTCTAG CCTCTGTAA GCTGGACTCG GCACGCCCGC AACAGCACC GAGGAGTTAA 840  
GAGAGCGCGA AGCGCAGGGA AGGCCTCCCT GCACGGGTGG GGGAAAGCGG CCGGTGCGAG 900  
GGCGGACAG GCACCTCGGC TGGCACTGGC TGCTAGGGAT GTCTCTCTGG ATAGGTGGG 960  
ATGACCCGCG CATGGCGCGG CTCTGGGGCT TCTGTGGCT GGTGTGGGC TTCTGGAGGG 1020  
CGCTTTCTGC CTGTCCCAAG TCCTGCAAAAT GCAGTGCCCTC TGGGATCTGG TGCAGCGACC 1080  
CTTCTCTGG CATCTGTGCA TTCTCCGAGT TGGAGCCCTAA CAGTGTAGAT CCTGAGAAC 1140  
TCACGAAAT TTTCATGCGA AACCGAATAA GGTTAGAAAT CATCAACGAA GATGATGTTG 1200  
AAGCTTATGT GGAAGTGAAG AATCTGACAA TTGTGGATTC TGGATTAAAA TTGTGGCTC 1260  
ATAAAGCATT TCTGAAAAAC AGCAACCTGC AGCAGATCAA TTTTACCGA AACAACTGA 1320  
CGAGTTTGTG TAGGAAACAT TTCCGTACCC TTGACTGTGC TGAAGTATC CTGGTGGGCA 1380  
ATCCATTTAC ATGCTCTGCT GACATATATG GGATCAAGAC TCTCCAAGAG GCTAAATCCA 1440  
GTCCAGACAC TCAGGATTGT TACTGCCTGA ATGAAAGCAG CAAGAATATT CCCCTGGCAA 1500  
ACCTCGAGAT ACCCAATTGT GGTTCGCCAT CTGCAAAATC GGCAGCAGCT AACCTCACTG 1560  
TGGAGGAAGG AAGTCTATC ACATTATCCT GTAGTGTGGC AGGTGATCG GTTCTAATA 1620  
TGTATTGGGA TGTGGTAAC CTGGTTTCCA AACATATGAA TGAACAAGC CACACACAG 1680  
GCTCCTTAAG GATAACTAAC ATTTTCATCG ATGACAGTGG GAAGCAGATC TCTGTGTGG 1740  
CGGAAATCTC TGATAGGAAA GATCAAGATT CTGTCAACCT CACTGTGCTA TTTCACCAA 1800  
CTATCAGATT TCTGAAATCT CCAACCTCAG ACCACCACTG GTGCATTCCA TTCACTGTGA 1860  
AAGCAACCCC AAGACAGCG CTTCAGTGGT TCTATAACGG GGCAATATG AATGAGTCCA 1920  
AATACATCTG TACTAAAATA CATGTTACCA ATCAGACGGA GTACCACGGC TGCTCCAGC 1980  
TGGTAATACC CACTCAGATG AACAATGGGG ACTACACTCT AATAGCCAGG AATGAGTATG 2040  
GGAAGGATGA GAAACAGATT TCTGCTCACT TCATGGGCTG GCCTGGAATT GACGATGGTG 2100  
CAAACCCAAA TTATCTGATG GTAATTTATG AAGATTATGG AACTGCAGCG AATGACATCG 2160  
GGGACACAC GAAAGAAATG AATGAAATCC CTTCACAGA CGTCACTGAT AAAACCGGTC 2220  
GGGAACATCT CTCGCTCTAT GCTGTGGTGG TGAATGCGTC TGTGGTGGGA TTTTGCTTT 2280  
TGGTAATGCT GTTCTGCTCT AAGTTGGCAA GACACTCCAA GTTTGGCATG AAAGGTTTTG 2340  
TTTGTGTTCA TAAGATCCCA CTGGATGGGT AGCTGAAATA AAGGAAAGGA CAGAGAAAGG 2400  
GGCTGTGGTG CTGTGTGGTT GATGCTGCCA TGTAAAGCTG ACTCTGGGGA CTGCTGTGG 2460  
CTTATCCCGG GAAGTCTGTC TTATCTGGGG TTTTCTGGTA GATGTGGGCG GTGTTGGAG 2520  
GCTGTACTAT ATGAAGCCTG CATATACTGT GAGCTGTGAT TGGGGAACAC CAATGCAGAG 2580  
GTAACCTCA GGCAGCTAAG CAGCACCTCA AGAAACATG TTAATTAAT GCTTCTCTTC 2640  
TTACAGTAGT TCAAAATACA AACTGAAATG AAATCCCATT GGATTGTACT TCTCTCTGA 2700  
AAAGTGTGCT TTTTACCCCT ACTGGACATT TATTGACTTA ATTGCTCTG TTTATTAAAA 2760  
TTGACCTGCA AAGTTAAAAA AAAATTAAG TTGAGAACAG GTATAAGTGC ACACGTAATA 2820  
GTCTAATCTA CATGTAACAC ATATTTTAGT GTGATTTTCT ATACTCTAAT CAGCACTGAA 2880  
TTCAGAGGGT TTGACTTTTT CATCTATAAC ACAGTGACTA AAAGAGTTAA GGGTATATAT 2940  
ACCATCACTT TGGGACTTGG TAGTATATTT AAAAGGTTAT TTCTTCACT GTCAATAAAA 3000  
GTCCAAATGT TTAGCTTAGG TCTGAGATC AAACAATGTT AAGGATTGTC TTAAAGTTCC 3060  
TTAGCCAGCA AAACAAAACA AAACAAAACA AAACAATGAA AAACGTTTAA AAAGAAGAAG 3120  
AAGAAAAAAA ACAAGAACAA GCAGCAACAG CTGTTTGTGT GGGGCTATAG ATTTAAGTTA 3180  
GGCAGTCA ATTTCAGAA AACTAAGAGT GGAATATATG CATATGGTGA AATTATAACC 3240  
TTGCCCTTTT TTATTTGCC CTCTGATGCC ACCTGCTTTT TAGAAGTCTG CCGAGTGAGA 3300  
AGCCACAGT ATCTCATGCT GTTGTGCTTA CAGAACTGCA GCTTTTCTAC TCTGAAAAGG 3360  
CTGGGAGCA GAATGGCTGG CCTGCTGTGA GCAGGAGAGG AGATTCTAAG AAGGATAGTC 3420  
CCCCCTACAA CACTCTGTCA TACTGCTGGG TTTTCATGGG TAGGAAAGCT TGTCTGACC 3480  
CCAGCAGCAA AGAGGTGCA GGTGCTAAT GAATATATGC TTTATAATGT CCTTCTCAT 3540  
TGCTGAGAGG GCAGCCTTAG AGCTGTGGAT TTCTGATCC CCCCTGAGTC TGAACCATGG 3600  
ACACCTGTTT CATTCACCTT AGCATCAGAG TGACCTTTGT ATGCTCTGTT CAGTCTGTGT 3660  
CAGGAGATAT GCTTGTCTCG AAGAGAGGTT TGGCTATCCC CACCCACACC CACCCACACC 3720  
TGTCTCTTTT TTATCAGAG GACTTCAGAG CCAGGCCATG AGCATTTTGT TTGAAAACAC 3780  
AATCAGCTCT GACAGTTAGA CATGCACACA GACGCCATAG CTGATTTGGA AACATTGATG 3840  
TTTTAAAAAT TTATTTTCTT TGGAAATAGT TGCACAAATG CTGCAATTTA GCTTTAAGGT 3900  
TCTATAGATT TTTAACTAGT CCAACACAGT CAGAACATT GTTTTGAATC CTCTGTAAAC 3960  
CAGGCAATTA ATCTTAATTA ACCAGGATCC ATTTAGGTAC CACTTGATAT AAAAAGGATA 4020  
TCCATAATGA ATATTTTATA CTGCATCCTT TACATTAGCC ACTAAATACG TTATTGCTGT 4080  
ATGAGAGACT TTCACAGAA CCTATGGATT GCAGCATTTT ACTTGGCTAC TTCATACCCA 4140



TGCCCTTAAAG AGGGGACAGTT TCTCAAAAGC AGAAACATGC CGCCAGTTCT CAAGTTTTC 4200  
 TCCTAACTCC ATTTGAATGT AAGGGCAGCT GGGCCCCAAT GTGGGGAGGT CCGAACATTT 4260  
 TCTGAATTC CATTTTCTGT TCGCGGCTA AATGACAGTT TCTGTCTATTA CTATGATTC 4320  
 GATCTTTCCC AAAGGTGTGT ATTTACAAAG AGGCCAGCTA ATAGCAGAAA TCATGACCTT 4380  
 GAAAGAGAGA TGAATTCAC GCTGTGAGCC AGGCCAGGAGC TCAGTATGGC AAAGGTTCTT 4440  
 GAGAATCAGC CATTTGGTGT AAAAAAGATT TTTAAAGCTT TTATGTATATA CCATGGAGCC 4500  
 ATAGAAAGGC TATGGATTGT TTAAGAACTA TTTTAAAGTG TTCCAGACCC AAAAAGGAAA 4560  
 AATAAAAAAA AAGGAATATT TGTACCCAAC AGCTAGAAGG ATTGCAAGGT AGATTTTTGT 4620  
 TTTAAATGG AGAGAAGTGG ACAGATAAGG CCATTTAATA TATCAAGAT CAGTTGACAT 4680  
 CTCCTAGGGA ATGATGAAAA CAGCAGGCTA T

Seq ID NO: 18 Protein sequence:  
 Protein Accession #: CAAS3571

1 11 21 31 41 51  
 MSSWIRNHGP AMARLWGFCH LVVGFWRAPF ACPTSCCKCSA SRIWCSDFSP GIVAFPRLEP 60  
 NSVDPEINTE IFIANQRRL EINEDVEAY VGLRNLITVD SGLKPVAKHA FLKNSNLQHI 120  
 NPTRNKLTSL SRKHFRLDL SELLLVGNPF TCSDIMWIK TLQEAQSSPD TQDLYCLNES 180  
 SKNIPLANLQ IFNCGLPAN LAAPNLVVEE GKSITLSCSV AGDPVPMYV DVGNLVSKEH 240  
 NETSHTQSL RITNISDDDS GKQISCAVEN LVGEDQDSVN LTVHFAPTIT FLESPTSDHE 300  
 WCIPPTVKGN PKPALQWFFN GAILNESKYI CTKIHVNET EYHGCLQLDN PTHMNGNDYT 360  
 LIAKNEYGKD EKQISAHFNG WPGIDDGANP NYPDVYEDY GTAANDIGDT TNRSNEIPST 420  
 DVTDKTGREH LSVYAVVVIA SVVGFCLLVM LFLKLARHS KFGMKGFVLF HKIPLDG

Seq ID NO: 19 DNA sequence  
 Nucleic Acid Accession #: NM\_000228  
 Coding sequence: 82-3600

1 11 21 31 41 51  
 GCTTTCAGGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTCTGG 60  
 GGATCACCCC ATTTGGCTGA GATGAGACCA TTCTTCTCTT TGTGTTTTCG CCTGCCTGGC 120  
 CTCCTGCATG CCCAACAGC CTGCTCCCGT GGGGCTGTCT ATCCACCTGT TGGGGACCTG 180  
 CTTGTGGGGA GGACCCGGTT TCTCCGAGCT TCATCTACCT GTGGACTGAC CAAGCCTGAG 240  
 ACCTACTGCA CCGAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCAGGCGAG 300  
 CCTCACAACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCTCCGG CCCCATGCGC 360  
 TGGTGGCAGT CCCAGAATGA TGTGAACCTT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420  
 TTCCAGCTTC AAGAAGTCAT GATGGAGTTC CAGGGGGCCA TGCCCGCCGG CATGCTGATT 480  
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540  
 AOCCTCACCT TCCTCGGGT CGGCCAGGGT CGGCCTCAGA GCTGGCAGGA TGTTCGCTGC 600  
 CAGTCCCTGC CTCAGAGGCC TAATGCACGC CTAATGGGG GGAAGGTCCA ACTTAACCTT 660  
 ATGGATTAG TGTCTGGAT TCCAGCAACT CAAAGTCAAA AAATTCAGA GGTGGGGGAG 720  
 ATCACAACT TGAGAGTCAA TTTCACCAGG CTGGCCCTGT TGCCCAAGG GGGCTACCAC 780  
 CCTCCAGAG CACTACTATG TGTGTCCAG CTCCGTCTGC AGGGGAGCTG CTTCTGTTCAC 840  
 GGCACTGCTG ATGCTGCTGC ACCCAAGCCT GGGGCTCTG CAGGCCCTCT CACCGCTGTG 900  
 CAGGTCCAGG ATGTCTGTGT CTGCCAGCAC AACACTCCGG GCCCAAATG TGAGCGCTGT 960  
 GCACCTTCT ACAAACAACG GCCCTGGAGA CGCGCGAGG GCCAGGACG CCATGAATGC 1020  
 CAAAGGTGG AACTGCAATG GCACTCAGAG ACATGTCACT TTGACCCCG TGTGTTTGGC 1080  
 CGCAGCGAG GGGCATATGG AGGTGTGTGT GACAATGGCC GGGACACAC CGAAGGCAAG 1140  
 AACTGTGAGC GGTGTGAGT GCACTATTTT CGGAACCGGC GCCCGGAGC TTCCATTGAG 1200  
 GAGACTGCTG TCTCTGCGA GTGTGATCCG GATGGGGCAG TGCCAGGGGC TCCCTGTGAC 1260  
 CCAGTGACG GGCAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320  
 AAGCCGGCT TCACTGAGT CACCTACGCC AACCGCAGG GCTGCCACG CTGTGACTGC 1380  
 AACATCTGG GGTCCCGGAG GGACATGCC TGTGACGAG AGAGTGGCG CTGCCCTTGT 1440  
 CTGCCCAAG CTGGCTTGAC CGGGCCCGC TGTGACCAGT GCCAGCGAG CTACTGCAAT 1500  
 AGTGGCCAGG GCTGTGAACC GTGTGCCCTG GACCCGACCA ACTCCCTCA GCCCAGATG 1560  
 CAACAGTTC CAGCTGTGG GCGCTGTGG GAAGGCTTG GTGGCTGTAT GTGCAGCGCT 1620  
 GCAGCCATCC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAG ATGCCGAGCC 1680  
 TGTGACTGTG ATTTCCGGGG AACAGAGGGC CCGGCTGTG ACAAGGCATC AGGCCGCTGC 1740  
 CTCTGCCGCG CTGGCTTGAC CGGGCCCGC TGTGACCAGT GCCAGCGAG CTACTGCAAT 1800  
 CGCTACCCCG TGTGGGTGGC CTGCCACCT TGTCTCAGA CCTATGATG GGACCTCCGG 1860  
 GAGCAGGCC TCGCTTTGG TAGACTCCG AATGCCACCG CCAGCCTGTG GTCAGGGCCT 1920  
 GGGCTGGAG ACCGTGGCT GGCCTCCCG ATCTAGATG CAAAGAGTAA GATTGAGCAG 1980  
 ATCCAGCAG TTCTCAGCAG CCCCGCAGT ACAGAGCAG AGGTGGCTCA GGTGGCCAGT 2040  
 GCCATCTCT CCTCAGGCG AACTCTCAG GGCCTGCAGC TGGATCTGCC CCTGGAGGAG 2100  
 GAGAGCTGT CCTTCCGAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCCTT 2160  
 ACTATGTATC AGAGGAAGAG GGAGCAGTTT GAAAAAATAA GCAGTGTCTGA TCCTTCAGGA 2220  
 GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGTCTA GCAGGTCTCC 2280  
 GACAGCTGCG GCCTTTTGGG CCACTCAGG GACAGCCGGA GAGAGGAGA GAGGCTGGTG 2340  
 CGGCAGGCG GAGGAGGAGG AGGCACCGGC AGCCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400  
 ATGTCTTGT TGCTTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCACGGCAG 2460  
 ATGGCTTGCA CCCCAATATC ATGCCCTGGT GAGCTATGTC CCCAAGACAA TGGCACAGCC 2520  
 TGTGGCTCCC GCTGCAGGGG TGTCTTCCC AGGGCCGGTG GGGCTTCTT GATGGCGGG 2580  
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCCAGCTCC AGCGGACCA GAGATGATT 2640  
 AGGGCAGCG AGGAATCTGC CTCACAGATT CAATCCAGTG CCCAGCGCTT GGAGACCCAG 2700  
 GTGAGCGCCA GCGCTCCCA GATGAGGAGG GATGTGAGAC GCACACGGCT CCTAATCCAG 2760  
 CAGGTGCGGG ACTTCCCTAAC AGACCCCGAC ACTGATGAG CCACTATCCA GGAGGTACAG 2820  
 GAGGCGCTGC TGGCCCTGTG GCTGCCACA GACTCAGCTA CTGTTCTGCA GAAGATGAAT 2880  
 GAGATCCAGG CCATTGACG CAGGCTCCCG AAGTGGACT TGGTGTCTGT CCAGACCAAG 2940  
 CAGGACATTG CGCGTGGCCG CCGGTTGCAG GCTGAGGCTG AGGAAGCCAG GAGCCGAGCC 3000  
 CATGCACTGG AGGGCCAGGT GGAAGATGTG GTTGGGAACC TGCGGCAGGG GACAGTGGCA 3060  
 CTCAGGAAG CTCAGGAAGC CATGCAAGGC ACCAGCGCT CCTTCCGCT TATCCAGGAC 3120  
 AGGGTGTCTG AGGTTCAGCA GGTACTCGCG CCAGCAGAAA AGCTGGTGAC AAGCATGACC 3180  
 AAGCAGCTGG GTGACTTCTG GACACGGATG GAGGAGCTCG GCCACCAAGC CCGGAGCAG 3240  
 GGGGAGAGG CAGTCCAGGC CCAGCAGCTT GCGGAAGGTG CCAGCAGCA GGCATTGAGT 3300  
 GCGCAAGAGG GATTGAGAG AATAAAACAA AAGTATGCTG AGTTGAAGGA CCGATTGGGT 3360

## WO 02/086443

CAGAGTTCCA TGCTGGGTGA GCAGSGTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420  
 GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480  
 CTGCGGGGCA GCCAGGCCAT CATGCTCGCG TCGCGGAGCC TGACAGGACT GGAGAAGCGT 3540  
 GTGGAGCAGA TCCGTGACCA CATCAATGGG CGCGTGCTCT ACTATGCCAC CTGCAAGTGA 3600  
 5 TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCGGCC TTTGCTTTTG GTTGGGGGCA 3660  
 GATTGGGTTG GAATGCTTTC CATCTCCAGG AGACTTTTAT GCAGCCTAAA GTACAGCCTG 3720  
 GACCACCCTT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAA 3780  
 GGGACAGTTA CACTTGACAG ACAAAGATGG TGGAGATTGG CATGCCATTG AAATCAAGAG 3840  
 10 CTCTCAAGTC AAGGAAGCTG GGCTGGGCGG TATCCCCCGC CTTTAGTTCT CCACTGGGGA 3900  
 GGAATCCTGG AACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960  
 AAAATCTTTG G

Seq ID NO: 20 Protein sequence:  
 Protein Accession #: NP\_000219

1 11 21 31 41 51  
 MRPFLLCPA LPGLLHAQQA CSRGACYPPV GDLLVGRTRF LRASSTOGLT KPETYCTQYG 60  
 EWQMKCKCD SRQPHNYSH RVENVASSSG PMRWQSQND VNPVSLQLDL DRRFQLQEVN 120  
 20 MEFGQMPAG MLIERSSDFG KTWVYQYLA ADCTSTFFRV RQGRPQSQND VRCOSLPQRP 180  
 NARLNGGKVQ LNLMDLVSGI PATQSQKIQE VGEITNLRVN FTRLAPVPQR GYHPPSAYYA 240  
 VSQRLQGGC FCHGHADRC A PKFGASAGPS TAVQVHDVCV CQHNTAGPNC ERCAPPYNNR 300  
 PWRPAEQDQA HEQQRCDNCG HSETCHFDPA VFAASQGAAY GVCNDCRDHT EGKNCERCQL 360  
 HYFNRPRFA SIQETCISCE CDPDGAVPGA PCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420  
 25 TYANPQGCHR CDNLGSRRR DMPCEDESGR CLCLPMVVPV KCDQCAPYHW KLASGQCEP 480  
 CADPHNSPQ PTVPVHRVA PCREGFGGLM CSAAAIRQCP DRTYGDVATG CRACDCFRG 540  
 TEGPGCDKAS GRCLCRPGLT GPRCQCQRG YCNRYFVCA CHPCFTYDA DLREQALRFG 600  
 RLNRNATASLW SGPGLDRGLT ASRILDAKSK IEQIRAVLSS PAVTEQVAVQ VASAILSLRR 660  
 30 TLQGLQLDLF LESETLSLFR DLESILDRSFN GLLTMYQRKR EQFEKISSAD PSGAFRMLST 720  
 AYEQSAQAAQ QVDSRRLLD QLRDSRREAE RLVRQAGGGG GTGSPKLVAL RLEMSLPLDL 780  
 TPTFNKLCGN SRQMACTPIS CPGELCPQDN GTACGSRERG VILPRAGAPL MAGQVAEQLR 840  
 GFNAQLQRTF QMIRAAEESA SQIQSSAQLR ETQVSASRSQ MEEDVRRTRL LIQVDRFLT 900  
 DEDTDAATIQ EVSEAVLALW LPTDSATVLQ KMNEIQATAA RLNVVDLVLS QTKQDIARAR 960  
 35 RLQAEAEAR SRHAHVEGVQ EDVVGHLRQG TVALQEAQDT MQGTSRSLRL IQDRVAEVQ 1020  
 VLRPAEKLVT SMTKQLGDFW TRMEELRHEQ RQQAARVQA QQLABGASEQ ALSAQEGFER 1080  
 1KQYABELKD RLQSSMLGE QGARIQSVKT EAEELPGETM EMDRMKOME LELLRGSQAI 1140  
 MLRSADLTGL EKRVQIRDH INGRVLYYAT CK

Seq ID NO: 21 DNA sequence  
 Nucleic Acid Accession #: NM\_003722  
 Coding sequence: 145-1491

1 11 21 31 41 51  
 45 TGTTGTATAT CAAAGACAGT TGAAGGAAAT GAATTTTGAA ACTTCACGGT GTGCCACCTT 60  
 ACAGTACTGC CCTGACCCCT ACATCCAGCG TTTCTGTAGAA ACCCAGCTCA TTCTCTTGG 120  
 AAAGAAAGTT ATTACCGATC CACCATGTCC CAGAGCACAC AGACAAATGA ATTCTCTAGT 180  
 CCAGAGGTTT TCCAGCATAT CTGGGATTTT CTGGAACAGC CTATATGTTC AGTTCAGCCC 240  
 50 ATTGAATTGA ACTTTGTGGA TGAACCATCA GAAGATGGTG CGACAAACAA GATTGAGATT 300  
 AGCATGGAAT GTATCCGCAT GCAGGACTCG GACCTGAGTG ACCCATGTGT GCCCAGTAC 360  
 ACGAACCTGG GCTCTCTGAA CAGCATGGAC CAGCAGATTC AGAACGGCTC CTGCTCCACC 420  
 AGTCCCTATA ACACAGACCA CGCGCAGAAC AGGTCACGGG CGCCCTCGCC CTACGCACAG 480  
 CCCAGCTCCA CCTTCGATGC TCTCTCTCCA TCACCCGCCA TCCCTCCCAA CACCGACTAC 540  
 CCAGGCCCGC ACAGTTTCGA CGTGTCTTC CAGCAGTGA GCACCCGCAA GTCCGCCACC 600  
 55 TGGACGTATT CCACTGAATC GAAGAACTC TACTGCCAAA TTGCARAGAC ATGCCCCATC 660  
 CAGATCAAGG TGATGACCCC ACCTCCTCAG GGAGCTGTTA TCCGCGCCAT CCCTGTCTAC 720  
 AAAAAGCTG AGCAGCTCAC GGAGGTGGTG AAGCGTGCC CCAACCATGA GCTGAGCCGT 780  
 GAATTCACG AGGGACAGAT TGCCCTCTCT AGTCATTGTA TTGAGTAGA GGGGAACAGC 840  
 CATGCCAGT ATGTAGAAGA TCCCATCACA GGAAGACAGA GTGTGCTGGT ACCTTATGAG 900  
 60 CCACCCCAAG TTGGCAGTGA ATTACAGACA GTCTGTGATA ATTTCTATGT TAACAGCAGT 960  
 TGTGTGGAG GGAATGAACG CGTCCCAATT TTAATCATTT TTAATCTGGA AACCAGAGAT 1020  
 GGGCAAGTCC TGGGCCGAGC CTGCTTTGAG GCCCGGATCT GTGCTTGCCC AGGAAGAGAC 1080  
 AGGAAGGCGG ATGAAGATAG CATCAGAAAG CAGCAAGTTT CGGACAGTAC AAAGAACGGT 1140  
 65 GATGGTACGA AGCCGCCGTT TGCTCAGAAC ACACATGGTA TCCAGATGAC ATCCATCAAG 1200  
 AAACGAAGAT CCCCAGATGA TGAAGTGTTA TACTTACCAG TGAGGGGCGG TGAGACTTAT 1260  
 GAAATGCTGT TGAAGATCAA AGAGTCCCTG GAATCATGCG GAATCATGCG TCAACACACA 1320  
 ATTGAACGCT ACAGGCAACA GCAACAGCAG CAGCACCAGC ACTTACTTCA GAAACATCTC 1380  
 CTTTACGCTT GCTTCAGGAA TGAGCTTGTG GAGCCCCGGA GAGAAACTCC AAAACAATCT 1440  
 70 GAGCTCTTCT TTAGACATTC CAAGCCCCCA AACCGATCAG TGTACCCATA GAGCCCTATC 1500  
 TCTATATTTT AAGTGTGTGT GTTGTATTTT CATGTGTATA TGTGAGTGTG TGTGTGTGTA 1560  
 TGTGTGTGCG TGTGTATCTA GCCCTCATAA ACAGGACTTG AAGACACTTT GGCTCAGAGA 1620  
 CCCAAGTCTT CAAAGGCACA AAGCCACTAG TGAGAGAATC TTTTGAAGGG ACTCAAACTT 1680  
 75 TTACAGAAA GGAATGTTTC TGCAGATTTT GTATCCTTAG ACCGGCCATT GGTGGGTGAG 1740  
 GAACCACTGT GTTGTCTCTG GAGCTTCTCT TTGTTTCTCT GAGGGGAGGG GTCAGGTGGG 1800  
 GAAAGGGGCA TTAAGATGTT TATTGGAACC CTTTCTGTCT TCTTCTGTCT GTTTTCTTAA 1860  
 AATTACAGG GAAGCTTTTG AGCAGGCTCT AAACTTAAGA TGTCTTTTAA AGAAAAGGAG 1920  
 AAAAAGTTTG TTAATGTCTG TGCATAAGTA AGTGTAGGT GACTGAGAGA CTCAGTCAGA 1980  
 CCCTTTAAAT GCTGGTCATG TAATAATATT GCAAGTAGTA AGAAACGAAG GTGTCAAGTG 2040  
 80 TACTGCTGGG CAGCGAGGTG ATCATTACCA AAGTAATCA ACTTGTGGG TGGAGAGTTC 2100  
 TTTGTGAGAA CTGTGATTAT TGTGTCTCTC CCCTCATGTG TAGGTAGAAC ATTTCTTAAT 2160  
 GCTGTGTACC TGCCCTCTGC ACTGTATGTT GGCATCTGTT ATGCTAAAGT TTTTCTGTGA 2220  
 CATGAAACCC TGAAGAGCCT ACTACAAAAA AACTGTGTGT TGGCCCCCAT AGCAGGTGAA 2280  
 CTCAATTTGT GCTTTTAAATA GAAAGACAAA TCCACCCAGG TAATATTGCC CTTACGTAGT 2340  
 85 TGTTTACCAT TATTCAAAGC TCAAAATAGA ATTGGAAGCC CTCTCACAAA ATCTGTGATT 2400  
 AATTGCTTAA ATTAGAGCTT CTATCCCTCA AGCCTACCTA CCATAAAACC AGCCATATTA 2460  
 CTGATACTGT TCAGTGCATT TAGCCAGGAG ACTTACGTTT TGAGTAAGTG AGATCCAAGC 2520  
 AGACGTGTTA AAATCAGCAC TCCTGGAGTG GAAATTAAG ATTGAAAGGG TAGACTACTT 2580

TTCTTTT TACTCAAAG TTTAGAGAAT CTCTGTTCT TTCCATTTA AAAACATATT 2640  
 TTAAGATAAT AGCATAAGA CTTTAAAAAT GTTCTCTCCC TCCATCTTCC CACACCCAGT 2700  
 CACCAGCACT GTATTTTCTG TCACCAAGAC AATGATTCTT TGTATTGAG GCTGTTGCTT 2760  
 TTGTGATGT GTGATTTTAA TTTTCAATAA ACTTTTGCACT CTGTTGTTAA AAGAAA

Seq ID NO: 22 Protein sequence:  
 Protein Accession #: NP\_003713

1 11 21 31 41 51  
 MSQSTQTNEF LSPEVPOHIW DFLEQPICSV QPIDLNFVDE PSEDGATNKI EISMDCIRMO 60  
 DSDLSDPMPW QYTNLGLLNS MDQIQNGSS STSPYNTDHA QNSVTAPSPY AQPSTTFDAL 120  
 SPSPAIPSNIT DYPGPHSFDV SFQSSSTAKS ATWYSTELK KLYCQIAKTC FIQIKVMTTP 180  
 PQGAVIRAMP VYKKAHVTE VVKRCFNHEL SREFNEGQIA PPSHLIRVEG NSHAQYVEDP 240  
 ITGRQSVLVP YEPQVGTET TTVLYNFMCN SSCVGGMNR R PILIIVTLET RDQVLGRRC 300  
 FEARICACPG RDRKADSDSI RKQVSDSTK NGDGTKRPR ONTHGIQMTS IKRRSPDDE 360  
 LLYLPVRGRE TYMLLKIKE SLELMQYLPO HTIETYROQO QQQHQLLQK HLLSACPRNE 420  
 LVEPRRETFK QSDVFFRHSK PPRRSVYP

Seq ID NO: 23 DNA sequence  
 Nucleic Acid Accession #: NM\_001944.1  
 Coding sequence: 84-3083

1 11 21 31 41 51  
 TTTTCTTAGA CATTAACTGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGACTT 60  
 TTTCAACAGG GAAATCAGAG ACAATGATGG GGTCTTCTCC CAGAATCACA GGGCTCTGG 120  
 CCATCTTCGT GGTGGTCATA TTGGTTCATG GAGAATTGCG AATAGAGACT AAAGGTCAT 180  
 ATGATGAAGA AGAGATGACT ATGCAACAGG CTAAGAAAGG GCAAAACGCT GAATGGGTGA 240  
 AATTTCGCAA ACCCTGCAGA GAAGGAGAGG ATAACCTCAA AAGAAACCCA ATTGCCAAGA 300  
 TTAATTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACCG AATCTCTGGA GTGGGAATCG 360  
 ATCAGCCGCC TTTTGGAAATC TTTGTTGTTG ACAAAACAC TGGAGATATT AACATAACAG 420  
 CTATAGTCGA CGGGAGGAAA ACTCCAAGCT TCCTGATCAC ATGTGCGGCT CTAATGCCCC 480  
 AAGGACTAGA TGTAGAGAAA CCATCTATAC TAACGGTTAA AATTTTGGAT ATTAATGATA 540  
 ATCTCCAGT AITTTCAAAA CAAATTTTCA TGGGTGAAAT TGAAGAAAAT AGTGCCCTCA 600  
 ACTCCTCGT GATGATACCT AATGCCACAG ATGCAGATGA ACCAAACCACT TTGAATTTCT 660  
 AAATTCGCTT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTTT CTCTAAGCA 720  
 GAAACACTGG GGAAGTCGCT ACTTTGACCA ATTCTCTTGA CCGAGAGCAA GCTAGCAGCT 780  
 ATCGTCTGTT TGTGAGTGGT GCAGACAAAG ATGGAGAGGG ACTATCAACT CAATGTGAAT 840  
 GTAATATTAA TGTGAAGAGT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900  
 CAGCAGGAT TGAAGAAAT ATTTTAAGTT CTGAATTTCT TCGATTTCAA GTAAACAGATT 960  
 TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATTT CTTTACCTCT GGGAAATGAAG 1020  
 GAAATGTTT TGAATACAA ACTGATCCTA GAACATATGA AGGCATCCTG AAATGGTGA 1080  
 AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAACCTTAG TATTGCTGTC AAAAACAAAG 1140  
 CTGAATTTCA CCAATCACTT ATCTCTCGAT ACCGAGTTCA GTCAACCCCA GTCACAAATC 1200  
 AGGTAATAAA TGTAGAGAAA GGAATTCAT TCCGTCTCGT TTCCAGACA TTTACTGTGC 1260  
 AAAAGGAGCT AAGTAGCAAA AAATGGTGG ATTATATCCT GGGAAACAT CAAGCCATCG 1320  
 ATGAGGACAC TAACAAAGCT GCCTCAAATG TCAATATGTT CATGGGAGCT AACGATGGTG 1380  
 GATACCTAAT GATTGATCA AAAACTGCTG AAATCAAAAT TGTCAAAAAT ATGAACCGAG 1440  
 ATTCTACTTT CATAGTTAAC AAAACAATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500  
 CGGGTAAAC TTCTACAGGC ACGGTATATG TTAGAGTACC CGATTTCAT GACAATTTGT 1560  
 CAACAGCTGT CCTCGAAAAA GATGCACTTT GCAGTTCTTC ACCTTCCGTG GTTGTCTCCG 1620  
 CTAGAACACT GAATATAGA TACACTGGCC CCTATACATT TGCACTGGAA GATCAACCTG 1680  
 TAAAGTTGCC TGCCGTATGG AGTATCACA CCCTCAATGC TACCTCGGCC CTCTCAGAG 1740  
 CCCAGGAACA GATACCTCTT GGAGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800  
 ACAATCGGTG TGAAGTGCCA CGCAGCTTGA CACTGGAAGT CTGTCACTGT GACAACAGGG 1860  
 GCATCTGTGG AACTTCTTAC CCAACCAACA GCGCTGGGAC CAGGTATGGC AGGCCGCACT 1920  
 CAGGAGAGCT GGGGCTGCCC GCCATCGGCC TGCTGCTCCT TGGTCTCCTG CTGCTGCTGT 1980  
 TGGCCCCCTT TCTGCTGTG ACCTGTGACT GTGGGCGAGG TTCTACTGGG GGAATGACAG 2040  
 GTGGTTTAT CCGAGTTCTT GATGGCTCAG AAGGAACAA TCACTAGTGG GGAATGAAG 2100  
 GAGCCCATCC TGAAGACAAG GAAATCACA ATATTGTGT GCCTCTCTGA ACAGCCAATG 2160  
 GAGCCGATTT CATGGAAGT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220  
 TGGAGGCAC TTCAAGATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280  
 GTGCTGACGG CTTTGCACA GGGACAGTGT CAGGAGCTGC TTCAGGATTG GGAGCAGCCA 2340  
 CTGGAGTTGG CATCTGTTC TCAGGGCAGT CTGGAACCAT GAGAACAAGG CATTCCACTG 2400  
 GAGGAACCAA TAAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTCTG GACTCCTACT 2460  
 TTTCTCAGAA AGCATTTGCC TGTGCGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520  
 TGTGATCTA TGATAATGAA GGCGCAGAT CCACTGGTTC TCCTGTGGGC TCCGTGGGTT 2580  
 GTTGAGTTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640  
 TTAATAAACT TGCAGAGATA AGCCTTGGTG TTGATGTTGA AGGCAAGAA GTTCAGCCAC 2700  
 CCTCTAAGA CAGCGGTAT GGGATTGAAT CCTGTGSCCA TCCATAGAA GTCCAGCAGA 2760  
 CAGGATTGT TAAGTGCCAG ACTTTGTGAG GAAGTCAAGG AGCTTCTGCT TTGTCGSCCT 2820  
 CTGGGTCTGT CCAGCCAGCT GTTCCATCC CTGACCCCTT GCAGCATGGT AACTATTAG 2880  
 TAAAGGAGAC TTAAGCGGCT TCTGTTTCCC TGTGCAACC TTCCACTGCA GGCTTTGATC 2940  
 CACTTCTCAC ACAAATGTG ATAGTGACAG AAAGGTGAT CTGTCCCAT TCCAGTGTTC 3000  
 CTGGCAACCT AGCTGGCCCA ACAGCAGTAC GAGGGTCACA TACTATGCTC GTACAGAGG 3060  
 ATCTTGTCTC TACTCTAATA TGACCAAGAT GAGCTGGAAT ACCACACTGA CCAATCTGG 3120  
 ATCTTGTGAC TAAAGTATTC AAAATAGCAT AGCAAGCTC ACTGTATTGG GCTAATATTT 3180  
 TGGCACTTAT TGACTTCTCT CATAACTGA TCACGATTAT AAATTAATG TTTGGGTTCA 3240  
 TACCCCAAAA GCATATGTT GTCACTCCTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA 3300  
 TCTTAAAGTT TTTCAAAACC CTAATATCAT ATTGCG

Seq ID NO: 24 Protein sequence:  
 Protein Accession #: NP\_001935.1

1 11 21 31 41 51

	MMGLFPRTTG	ALAIFVVVIL	VHGLRIETK	QYDEEEMTM	QAKRRQKRE	WVKFAKPCRE	60
	GEDNSKRNPI	AKITSDYQAT	QKITRYISGV	GIDQPPFGIF	VVDKNTGDI	ITAIVDREET	120
5	PSPLITCRAL	NAQGLDVEKP	LILTVKILDI	NDNPPVFSQ	IFMGBIEENS	ASNSLVMLN	180
	ATDADEPNEL	NSKIAFKIVS	QEPAGTFMFL	LSRNTGEVTR	LTSNLDREQA	SSVRLVVSQA	240
	DKDGEGLSTQ	CECNIKVKDV	NDNPPMFRDS	QYSARIEENI	LSSELLRFQV	TDLDEEYTDN	300
	WLAVYFPTSG	NEGNWFETQT	DPRTNEGILK	VVKALDYBQL	QSVKLSIAVK	NKAEFHQSIV	360
	SRVRQSTFV	TIQVINVRG	IAPRPASKTF	TVQKGISSKK	LVDYILGTQY	AIEDTNKAA	420
10	SNVKYVMGRN	DGYVIMDSK	TAEIKFVKNM	NRDSTFIVNK	TITAEVLAI	EYTKTSTGT	480
	VYVRVPDFND	NCPTAVLEKD	AVCSSSPSVV	VSARTLNNRY	TGPYTFALD	QPVKLEAVNS	540
	ITTLNATSL	LRAQEQIPPG	VYHISLVLT	SQNNRCCEMR	SLTILEVQCD	NRGICGTSYP	600
	TTSPGTRYGR	PHSGRLGPAA	IGLLLLGLLL	LLAPLILLT	CDGAGSTGG	VTGGFIFVPD	660
	GSEGTIHQWG	IEGAHPEDKE	ITNICVFPVT	ANGADFMESS	EVCTNTYARG	TAVEGTSME	720
15	MTTLGAATE	SGGAAGFATG	TVSGAASGFG	AATGVGICSS	QSGGTMTRH	STGGTNKDYA	780
	DGAISLWFLD	SYFSQKAFAC	AEEDDQGEAN	DCLLIYDNEG	ADATGSPVGS	VGCCSFIADD	840
	LDDSFLLSLG	PKFKKLARIS	LGVDGEGKEV	QPPSKDSGYG	IESCGHPIEV	QQTGFVKCQT	900
	LSGQSGASAL	SAGSVQPAV	SIPDPLQHN	YLVETYSAS	GSVLQPSSTAG	FDPLLTQNV	960
	VTERTICPIS	SVPGNLAGPT	QLRGSHTMLC	TEDPCSRLL			

Seq ID NO: 25 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 56-1642

25	1	11	21	31	41	51	
	AGTATCCCAG	GAGGAGCAAG	TGGCACCTCT	TCGGACCTAG	GCTGCCCTGT	CCGTCACTGC	60
	GCAAGGGATC	CTTCTCCCG	CAGCGGGCTT	GCTGTCCGAT	GACGATGTCG	TAGTTTCTCC	120
	CATGTTTQAG	TCCACAGCTG	CAGATTTCGG	GTCTGTGGTA	CGCAAGAAC	TGCTATCAGA	180
30	CTGCTCTGTC	GTCTCTACCT	CCCTAGAGGA	CAAGCAGCAG	GTTCATCTG	AGGACAGTAT	240
	GGAGAGGCTG	AAAGTATATC	TGAGGGTTAG	GCCCTTGTTA	CCTTCAGAGT	TGGAACGACA	300
	GGAAGATCAG	GGTTGTGTCC	GTATTGAGAA	TGTGGAGACC	CTTGTTCTAC	AAGCACCCAA	360
	GGACTCTTTT	GCCCTGAAGA	GCAATGAACG	GGGAATTGGC	CAAGCCACAC	ACAGGTTTAC	420
	CTTTTCCCAG	ATCTTTGGGC	CAGAAGTGGG	ACAGGCATCC	TTCTTCAACC	TAAGTGTGAA	480
35	GGAGATGGTA	AAGGATGTAC	TCAAAGGGCA	GAATCTGGCT	ATCTATACAT	ATGGAGTCAC	540
	TAACCTCAGG	AAAACCCACA	CGATTCAAGG	TACCATCAAG	GATGGAGGGA	TTCTCCCCCG	600
	GTCCCTGGCG	CTGATCTTCA	ATAGCCTCCA	AGGCCAATCT	CATCCACAC	CTGATCTGAA	660
	GCCCTTGCTC	TCCAATGAGG	TAATCTGGCT	AGACAGCAAG	CAGATCCGAC	AGGAGGAAAT	720
	GAAGAAGCTG	TCCCTGCTAA	ATGGAGGCCT	CCAAGAGGAG	GAGCTGTCCA	CTTCTTGA	780
40	GAGGAGTGTC	TACATCGAAA	GTGGATAGG	TACCAGCACC	AGCTTGCACA	GTGGCATTGC	840
	TGGGCTCTCT	TCTATCAGTC	AGTGTACCA	CAGTAGCCAG	CTGGATGAAA	CAAGTCATCG	900
	ATGGGCACAG	CAGACACACT	CCCCACTACC	TGTCGCCGCA	AACATTCGCT	TCTCCATCTG	960
	GATCTCATTC	TTTGAGATCT	ACAACGAATC	GCTTTATGAC	CTATTAGAAC	CGCCTAGCCA	1020
	ACAGCGCAAG	AGGCAGACTT	TGCGGCTATG	CGAGGATCAA	AATGGCAATC	CCTATGTGAA	1080
45	AGATCTCAAC	TGGATTCAATG	TGCAAGATGC	TGAGGAGGCC	TGGAAGCTCC	TAAAGTGGG	1140
	TCGTAAAGAC	CAGAGCTTTG	CCAGCACCCA	CCTCAACCA	AATCCAGCC	GCAGTCACAG	1200
	CATCTTCTCA	ATCAGGATCT	TACACCTTCA	GGGGGAAGGA	GATATAGTCC	CAAGATCAG	1260
	CGAGCTGTCA	CTCTGTGATC	TGGCTGGCTC	AGAGCGCTGC	AAAGATCAGA	AGAGTGGTGA	1320
	ACGGTTGAAG	GAAGCAGGAA	ACATTAAAC	CTCTCTACAC	ACCCTGGGCC	GCTGTATTGC	1380
50	TGCCCTTCGT	CAAAACCAGC	AGAACCAGTC	AAAGCAGAAC	CTGGTTCCCT	TCCGTGACAG	1440
	CAAGTTGACT	CGAGTGTTC	AAGGTTTCTT	CACAGGCCGA	GGCGTTCTCT	GCATGATTGT	1500
	CAATGTGAAT	CCTGTGCAAT	CTACCTATGA	TGAAACTCTT	CATGTGGCCA	AGTTCTCAGC	1560
	CATTGTGATC	CAGGTGACTT	GTGCATGCC	CACCTATGCA	ACTGGGATTC	CCATCCCTGC	1620
	ACTCGTTCAT	CAAGGAACAT	AGTCTTCAGG	TATCCCCCAG	CTTAGAGAAA	GGGGCTAAGG	1680
55	CAGACACAGG	CCTTGATGAT	GATATTGAAA	ATGAAGCTGA	CATCTCCATG	TATGGCAAAG	1740
	AGGAGCTCCT	ACAAGTTTGT	GAAGCCATGA	AGACACTGCT	TTTGAAGGAA	CGACAGGAAA	1800
	AGCTACAGCT	AGGATGCTG	CTCCGAGATG	AAATTTGCAA	TGAGATGGTA	GAACAGATGC	1860
	AACAGCGGGA	ACAGTGTGTC	AGTGAACATT	TGGACACCCA	AAAGGAAGTA	TTGGAGGAAA	1920
	TGTATGAAGA	AAAACCTAAT	ATCCTCAAGG	AGTCACTGAC	AAGTTTTTAC	CAAGAAGAGA	1980
60	TTCAGGAGCG	GGATGAAAAG	ATTGAAGAGC	TAGAAGCTCT	CTTGCAGGAA	GCCAGACAAC	2040
	AGTCAGTGGC	CCATCAGCAA	TCAGGGTCTG	AATTGGCCCT	ACGGCGGTCA	CAAAGGTTGG	2100
	CAGCTTCTGC	CTCCACCCAG	CAGCTTCAGG	AGGTTAAAGC	TAAATTACAG	CAGTGCAAGG	2160
	CAGAGCTAAA	CTCTACCACT	GAAGAGTTGC	ATAAGTATCA	GAAATGTGTA	GAACCAACAC	2220
	CCTCAGCCAA	GCCCTTCACC	ATTGATGTGG	ACAAGAAGTT	AGAAGAGGGC	CAGAAGAATA	2280
65	TAAGGCTGTT	GCGACAGAG	CTTCAGAAAC	TTGGTGAGTC	TCTCCAATCA	GCAGAGAGAG	2340
	CTTGTTGCCA	CAGCACTGGG	GCAGGAAAAC	TTCTGTCAGC	CTTCACCACT	TGTGATGACA	2400
	TCTTAATCAA	ACAGGACCAG	ACTCTGGCTG	AACGACAGAA	CAACATGGTG	CTAGTGAAAC	2460
	TGGACCTTGG	GAAGAAGGCA	GCATGTATTG	CTGAGCAGTA	TCATATCTGT	TTGAACTCC	2520
	AAGGCCAGTG	TTCTGCCAAA	AAGCGCCTTG	GTACCAACCA	GGAAATCAG	CAACCAAAAC	2580
70	AACAAACACC	AGGGAAGAAA	CCATTCTTTC	GAAATTTACT	TCCCGGAACA	CCAACCTGCC	2640
	AAAGCTCAAC	AGACTGCAGC	CCTTATGCC	GGATCCTACG	CTCACGGCGT	TCCCTTTTAC	2700
	TCAAATCTGG	GCCCTTTTGG	AAAAGTACT	AAGGCTGTGG	GGAAGAGAAA	GAGCAGTCAT	2760
	GGCCCTGAGG	TGGGTGAGCT	ACTCTCTGTA	AGAAATAGGT	CTCTTTTATG	CTTTACCATTA	2820
	TATCAGGAAT	TATATCCAGG	ATGCAATACT	CAGACACTAG	CTTTTCTCTC	ACTTTTGTAT	2880
75	TATAACCAAC	TATGTAATCT	CATGTGTTG	TTTTTTTTTA	TTTACTTATA	TGATTCTTAT	2940
	GCACACAAA	ACAGTTATAT	TAAAGATATT	ATTGTTTACA	TTTTTTTATG	AATTCCAAAT	3000
	GTAGCAAAAT	CATTAAACA	AATTATAAAA	GGGACAGAAA	AA		

Seq ID NO: 26 Protein sequence:  
Protein Accession #: Eos sequence

80	1	11	21	31	41	51	
	MSQGLSPPA	GLLSDDVVV	SPMFESTAAD	LGSVVRKNLL	SDCSVVSTSL	EDKQVPSSE	60
	SMKVKVYLR	VRPLLPSELE	RQEDQGCVR	ENVETLVLA	PKDSFALKSN	ERGIGQATER	120
85	FTFSQIFGPE	VQGSFNLPT	VKEMVKDVLK	QONWLIYTYG	VINSKTHTI	QGTIKDGGIL	180
	PRSLALIFNS	LQGLHPTPD	LKPLLSNEVI	WLDKQLRQE	EMKLSLLNG	GLQEEELSTS	240
	LKRSVYIEBR	IGTSTSPDSG	IAGLSSISQC	TSSSQLDETS	HRWAQPDATP	LPVPANIRFS	300

IWISFPEIYN ELLYDLLEPP SQQRKRQTLR LCEQDQNGNPF VKDLNWIHVQ DAEAEWKLLK 360  
 VGRKNQSPAS THLNQNSSRS HSIFSRILH LQEGGDIVFK ISELSLCDLA GSERCKDQKS 420  
 GERLKEAGNI NTSLHTLGRC IAA LRQNNQN RSKQNLVPPFR DSKLTVFQGG FFTGRGRSCM 480  
 IVNVNFCAST YDSTLELVAKF SAIASQVTC APTATGIPI PALVHQGT

Seq ID NO: 27 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 13-1424

1 11 21 31 41 51  
 TAGAAGTTTA CAATGAAGTT TCTTCTAATA CTGCTCCTGC AGGCCACTGC TTCTGGAGCT 60  
 CTTCCCCTGA ACAGCTCTAC AAGCCTGGAA AAAATAATG TGCTATTG TGAAAGATAC 120  
 TTAGAAAAAT TTTATGGCCT TGAGATAAAC AAACCTCCAG TGACAAAAAT GAAATATAGT 180  
 GGAAACTTAA TGAAGGAAAA AATCCAAGAA ATGCAGCACT TCTTGGGTCT GAAAGTGACC 240  
 GGGCAACTGG ACACATCTAC CCTGGAGATG ATGCACGCAC CTCGATGTGG AGTCCCGCAT 300  
 GTCCATCATT TCAGGGAAT GCCAGGGGG CCGTATGGA GGAACATTA TATCACCTAC 360  
 AGAATCAATA ATTACACACC TGACATGAAC CGTGAGGATG TTGACTAGC AATCCGGAAA 420  
 GCTTTCCAAAG TATGGAGTAA TGTACCCCTC TTGAAATCA GCAAGATTAA CACAGGCATG 480  
 GCTGACATT TGTGGTATT TGCCCGTGG GCTCATGGAG ACTTCCATGC TTTTGATGGC 540  
 AAAGGTGGAA TCCTAGCCCA TGCTTTTGA CCTGGATCTG GCATTGGAGG GGATGCACAT 600  
 TTCGATGAGG ACGAATTCTG GACTACACAT TCAGGAGGCA CAACTTGT CTCTACTGCT 660  
 GTTCAAGAGA TTGGCCATTC CTTAGGTCTT GGCATTCTA GTGATCCAAA GGCCGTAAATG 720  
 TTCCCACTCT ACAATATATG TGACATCAAC ACATTTCGCC TCTCTGCTGA TGACATACGT 780  
 GGCATTCACT CCTGTATGAG AGACCCAAAA GAGAACCAC GCTTGCCAAA TCTGACAAAT 840  
 TCAGAACCAAG TCTCTGCTG CCCCATTG AGTTTGTATG CTGTCACTAC CGTGGGAAAT 900  
 AAGATCTTTT TCTTCAAGA CAGGTCTCTC TGGCTGAAGG TTCTGAGAG ACCAAAGACC 960  
 AGTGTAAATT TAATTTCTTC CTTAGGCCCA ACCTTGCCAT CTGCATTGA AGCTGCTTAT 1020  
 GAAATTGAAG CCAGAAATCA AGTTTTCTT TTTAAAGATG ACRAATAGT GTTAATTAGC 1080  
 AATTTAAGAC CAGAGCCAAA TTATCCCAAG AGCATACATT CTTTTGGTTT TCTTAACATT 1140  
 GTGAAAAAAA TGTATGCACT TGTTTTTAA CCAAGTTTTT ATAGGACCTA CTCTTTGTA 1200  
 GATAACCACT ATTGAGGATA TGATGAAAGG AGACAGATGA TGGACCTCGG TTATCCCAAA 1260  
 CTGATTACCA AGAATCTCCA AGGAATCGGG CCTAAAATTG ATGCAGTCTT CTACTCTAAA 1320  
 AACAAATACT ACTATTCTT CCAAGGATCT AACCAATTG AATATGACTT CTTACTCCAA 1380  
 CGTATCACCA AAACACTGAA AAGCAATAGC TGGTTTGGTT GTTGAAGATG GTGTAATTAA 1440  
 TGGTTTGGT TAGTTCAGCT CAGCTTAATA AGTATTATG GCATATTGTC TATGTCTCTA 1500  
 GTGTACCACT ACTTAGAGAT ATGTATCATA AAAATAAAAT CTGTAAACCA TAGGTAAATG 1560  
 TTATATAAAA TACATAAATAT TTTTCAATT TGAAACTCT AATTGTCCAT TCTTGCTTGA 1620  
 CTCTACTATT AAGTTTGAAA ATAGTTACCT TCAAAGCAAG ATAATTCTAT TTGAAGCATG 1680  
 CTCCTGTAAG TGCTTCCTAA CATCCTTGGA CTGAGAAATT ATACTTACTT CTGGCATAAC 1740  
 TAAATTAAG TATATATATT TTGGCTCAAA TAAATTTG

Seq ID NO: 28 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MKFLILLILQ ATASGALPLN SSTSLKNNV LFGERYLEKF YGLEINKLPV TKMKYSGNLM 60  
 KEKIQEMQHF LGLKVTGQLD TSTLEMHAP RCGVPDVHFF REMPGGPVWR KHYITYRINN 120  
 YTPDMNREDV DVAIRKARQV WSNVTPLKPS KINTGMADIL VVFARGAHDG FHAFDGRKGI 180  
 LAHAFGPQSG IGDADHPEDF EPWTHSGGT NLFLTAVHEI GHSGLGLHSS DPKAVMPFTY 240  
 KYVDINTFRL SADDIRGQS LYGDPKENQR LPNPDNSEPA LCDPNLSFDA VTTVGNKIFP 300  
 FKDRFPWLK SERPFTSVNL ISSLWPTLPS GIEAAYEIEA RNQVFLFKDD KYWLISNLRP 360  
 EPNYKSIHS FGFNPFVKKI DAAVFNPRFY RTVPVDNQY WRYDERRQMM DPGVPEKLITK 420  
 NFQGIQPKID AVFYSKNKY YFFQGSNQFE YDFLLQRIK TLKSNWFGC

Seq ID NO: 29 DNA sequence  
 Nucleic Acid Accession #: NM\_006115.1  
 Coding sequence: 236..1765

1 11 21 31 41 51  
 GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGGCT GCAGCCCTC AGCACCGCTC 60  
 CGGGACACCC CACCCGCTTC CCAGGCGTGA CCTGTCAACA GCAACTTCGC GGTGTGGTGA 120  
 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180  
 GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC CTAAGTCGCT TCAAAATGGA 240  
 ACGAAGGCGT TTGTGGGGTT CCATTTCAGAG CCGATACATC AGCATGAGTG TGTGACAAG 300  
 CCCACGAGGA CTTGTGGAGC TGGCAGGGCA GAGCCTGCTG AAGGATGAGG CCTGGCCAT 360  
 TGCCGCCCTG GAGTTGCTGC CCAGGGAGCT CTTCGCCCA CTCTTCATGG CAGCCTTTGA 420  
 CGGGAGACAC AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CTTGCTTCCC 480  
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540  
 TGGACTTGAT GTGCTCTCTG CCCAGGAGGT TCGCCCCAGG AGGTGGAAC TTCAGTGCT 600  
 GGATTTCAGG AAGAACTCTC ATCAGGACTT CTGACTGTA TGGTCTGGAA ACAGGGCCAG 660  
 TCTGTACTCA TTTCCAGAGC CAGAAGCAGC TCAGCCCATG ACAGAGAGC GAAAGTAGA 720  
 TGGTTTGGAG ACAGAGGCAG AGCAGCCCTT CATTCAGTA GAGGTGCTCG TAGACTGT 780  
 CCTCAAGGAA GGTGCTCTG ATGAATTGTT CTCTACCTC ATTGAGAAAG TGAACGGA 840  
 GAAAAATGTA CTACGCTGCT GCTGTAAGAA GCTGAAGATT TTTGCAATGC CCATGCAGGA 900  
 TATCAAGATG ATCCTGAAAA TGGTGACGCT GGACTCTATT GAAGATTGAG AAGTGACTTG 960  
 TACCTGGAGG CTACCCACTT TGGCGAAATT TTCTCTTAC CTGGGCCAGA TGATTAACT 1020  
 GGTATGACTC CTCTCTCCC ACATCCATGC ATCTTCTTAC ATTTCCCGGC AGAAGGAAGA 1080  
 GCAGTATATC GCCCAGTTCA CTCTCAGTT CTCTCAGTCT CAGTGCCTGC AGGCTCTCTA 1140  
 TTGGGACTCT TTAATTTTTC TTAGAGGCCG CTGGGATCAG TTGCTCAGGC ACCTGATGAA 1200  
 CCCCTTGGAA ACCCTCTCAA TAACCTAAGT CCGGCTTTG GAAGGGGATG TGATGCATCT 1260  
 GTCCAGAGT CCCAGCGTCA GTCAAGCTAAG TGTCTGAGT CTAAGTGGGG TCATGCTGAC 1320  
 CGATGTAAGT CCGAGGCCCT TCCAAGCTCT GCTGGAGAGA GCCTCTGCCA CCTCCAGGA 1380  
 CCTGGTCTTT GATGAGTGTG GATCAAGGGA TGATCAGCTC CTGCGCTTCC TGCTCTCTCT 1440  
 GAGCCACTGC TCCAGCTTAA CAACCTTAAG CTCTACGGG AATTCATCTC CCATATCTGC 1500

CTTGACAGT CTCCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560  
 TGTCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCTC CACCTGGAGA GGCTTGCCTA 1620  
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCACGCA TGGTCTGGCT 1680  
 TAGTGCCAC ACCTGTCTCT ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740  
 GTGCCCTGT TTCATGCCCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800  
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTTGAAGCAA CAAAGCAGCC ACAGTTTCAG 1860  
 ACAAAATGTT AGTTGTAGTG AGGAAAAACAT GTTCAGTGAG GAAAAAACAT TCAGACAAAT 1920  
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGACTTG AGGAGTTAAT 1980  
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040  
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100  
 TGTGAAAT AAAGAGAAGC AATGTGAAGC AAAAAAAAAA AAAAAAAA

Seq ID NO: 30 Protein sequence:  
 Protein Accession #: NP\_006106.1

1 11 21 31 41 51  
 GCTTCAGGGT ACAGCTCCCC CGCAGCCAGA AGCCGGGCTT GCAGCGCCTC AGCACCCTC 60  
 CGGGACACCC CACCCGCTTC CCAGGCGTGA CCGTCTCAAC GCAACTTCGC GGTGTGGTGA 120  
 ACTCTCTGAG GAAAAACCAT TTTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180  
 GAGACCTAGA AATCCAAGCG TTGGAGGTCC TGAGGCCAGC CTAAGTGCCT TCAAAATGGA 240  
 ACGAAGGCGT TTGTGGGTTT CCATTTCAGAG CGATACATC AGCATGAGTG TGTGGACAAG 300  
 CCCACGAGGA CTGTGTGAGC TGGCAGGCGA GAGCCTGCTG AAGGATGAGG CCCTGGCCAT 360  
 TGCCGCTCTG GAGTTGCTGC CCAGGAGGCT CTTCCCGCCA CTCTTCATGG CAGCCTTTGA 420  
 CGGGAGACAC AGCCAGACCC TGAAGGCAAT GGTGCAGGCC TGGCCCTTCA CCTGCTCTCC 480  
 TCTGGGAGTG CTGATGAAGG GACAACTCT TCACCTGGAG ACCTTCAAAG CTGTGCTTGA 540  
 TGGACTTGAT GTGCTCTTGT CCCAGGAGGT TCGCCCCAGG AGTGGAAAC TTCAGTGTCT 600  
 GGATTACGCG AAGAATCTCT ATCAGGACTT CTGGACTGTA TGGTCTGGAA ACAGGGCCAG 660  
 TCTGTACTCA TTTCCAGAGC CAGAAGCAGC TCAGCCCATG ACAAGAAGAG GAAAAAGTGA 720  
 TGGTTTGAAG ACAGAGGCGG AGCAGCCCTT CATTCCAGTA GAGGTGCTCG TAGACCTGTT 780  
 CCTCAAGGAA GGTGCTGTG ATGAATTGTT CTCTACCTC ATTGAGAAAG TGAAGCGAAA 840  
 GAAAAATGTA CTACGCTGTG GCTGTGAAG GCTGAAGATT TTTGCAATGC CCATGCAGGA 900  
 TATCAAGATG ATCCTGAAAA TGGTGAGCT GGACTCTATT GAAGATTGAG AAGTGACTTG 960  
 TACCTGGGAG TACCCACCTT TGGCGAAATT TTCTCTTAC CTGGGCCAGA TGATTAACTC 1020  
 GCGTAGACTC CTCTCTCTCC ACATCCATGC ATCTTCCCTAC ATTTCCTCCG AGAAGGAAGA 1080  
 CAGATATATC GCCAGTTCA CCTCTAGTT CCTCAGTCTG CAGTGCCTGC AGGCTCTCTA 1140  
 TGTGACTCTT TTAATTTTTC TTAGAGGCGG CCTGGATCAG TTGCTCAGGC ACGTGATGAA 1200  
 CCCCTTGGAA ACCCTCTCAA TAACTAACTG COGGCTTTCG GAAGGGGATG TGATGCATCT 1260  
 GTCCAGAGAT CCCAGCGTCA GTCAGCTAAG TGTCTGAGT CTAAGTGGGG TCATGCTGAC 1320  
 CGATGTAAGT CCCAGCCCCC TCCAAGCTCT GCTGAGAGA GCCTCTGCCA CCCTCCAGGA 1380  
 CTTGGTCTTT GATGAGTGTG GGATCAAGGA TGATCAGCTC CTTGCCCTCC TGCTCTCTCT 1440  
 GAGCCACTGC TCCAGCTTCA CAACTTAAG CTCTACGGG AATTCATCT CCATATCTGC 1500  
 CTTCGAGAGT CTCCTGCAGC ACCTCATCGG GCTGAGCAAT CTGACCCACG TGCTGTATCC 1560  
 TGTCCCCTG GAGAGTTATG AGGACATCCA TGGTACCCTC CACCTGGAGA GGCTTGCCTA 1620  
 TCTGCATGCC AGGCTCAGGG AGTTGCTGTG TGAGTTGGGG CGGCCACGCA TGGTCTGGCT 1680  
 TAGTGCCAC ACCTGTCTCT ACTGTGGGGA CAGAACCTTC TATGACCCGG AGCCCATCCT 1740  
 GTGCCCTGT TTCTGCTCTA ACTAGCTGGG TGCACATATC AAATGCTTCA TTCTGCATAC 1800  
 TTGGACACTA AAGCCAGGAT GTGCATGCAT CTTGAAGCAA CAAAGCAGCC ACAGTTTCAG 1860  
 ACAAAATGTT AGTTGTAGTG AGGAAAAACAT GTTCAGTGAG GAAAAAACAT TCAGACAAAT 1920  
 GTTCAGTGAG GAAAAAAGG GGAAGTTGGG GATAGGCAGA TGTGACTTG AGGAGTTAAT 1980  
 GTGATCTTTG GGGAGATACA TCTTATAGAG TTAGAAATAG AATCTGAATT TCTAAAGGGA 2040  
 GATTCTGGCT TGGGAAGTAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAAGAAAC 2100  
 TGTGAAAT AAAGAGAAGC AATGTGAAGC AAAAAAAAAA AAAAAAAA

Seq ID NO: 31 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 64-2754

60 1 11 21 31 41 51  
 GGCAGGTCTC GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60  
 CCGATGGCGG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120  
 CTGACCCCTG TGATCTTCAG TCGTGTGCTG GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180  
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240  
 TCTGCAGACC TCATCCGCTC AAGTGATCCT GATTTTCAGAG TTCTAAATGA TGGGTCAGTG 300  
 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360  
 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420  
 TCGAAGACAA GACACACTAG AGAACTGTT CTGAGCGGTG CCAAGAGGAG ATGGGCACCT 480  
 ATTCTTGTCT CTATGAAGA GAATTCCTTG GGCCTTTTCC CATTGTTTCT TCAACAAGTT 540  
 GAATCTGATG CAGCAGAGAA CTATCTGTC TTCTACTCAA TAAGTGGAGC TGGAGTTGAT 600  
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGACTCGG 660  
 CCTGTGGATC GTGAAGAATA TGATGTTTTT GATTTGATTG CTTATGCGTC AACTGCAGAT 720  
 GGATATTCTG CAGATCTGCC CCTCCACTA CCCATCAGGG TAGAGGATGA AAATGACAAAC 780  
 CACCCCTGTT TCACAGAAGC AATTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840  
 ACTACAGTGG GGTGTTTGTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACCGGC 900  
 CTGAAATACA GCATTTTGCA CGACACACCA AGGTCACTTG GGCTCTTTTC TGTGCATCCC 960  
 AGCAGGCGG TAATCACAC AGTCTCTCAT TATTTGGACA GAGAGGTTGT AGACAAGTAC 1020  
 TCATTGATAA TGAAGTACA AGACATGGAT GGCAGTTT TTGGATTGAT AGGCACATCA 1080  
 ACTTGATACA TAACGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAATGCT 1140  
 TATGAAGCAT TTGTAGAGGA AAATGCATTG AATGTGAAA TCTTACGAAT ACCTATAGAA 1200  
 GATAAGGATT TAATTAAAC TGCCAATTGG AGAGTCAATT TTACCATTTT AAAGGGAAT 1260  
 GAAAAATGAC ATTTCAAAAT CAGCAGAGAC AAGAAACTA ATGAAGGTG TCTTCTGTT 1320  
 GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGAACC TGAATTTGG AGTAAACAAT 1380  
 GAAGCCCAT TTGCTAGAGA TATTCACAGA GTGACAGCCT TGAACAGAGC CTTGTTTACA 1440  
 GTTCATGTGA GGGATCTGGA TGAGGGGCTT GAATGCACTC CTGACGCCA ATATGTGCGG 1500  
 ATTAAAGAAA ACTTAGCAGT GGGGTCAAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560

	AAATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCGATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCCAAAATG	AGTTGTATATA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
5	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAATGATA	ATCCACCAGA	AATACTTCAA	1800
	GAATATGTAG	TCATTGTCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTGAGTTTTC	CCAATCTTTC	TCCGAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAAATTAAAT	GATACAGCTG	CCCCTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAAATA	TACCATTCTT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
10	GCAACAAAAT	TATGTAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACTCA	GTGCTGTGGG	2100
	ACTTCAAGGA	GTACAGGAGT	AATACTTGGG	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCCTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTTTGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCTGTGAAGA	TTTAGCACAG	CAAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAAACTAC	CAACAACCTCT	2340
15	AGCCAGAGGT	TTTGTGGTAC	TATGGGATCA	GGAATGAAAA	ATGGAGGGCA	GGAAACCAAT	2400
	GAATATGATG	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCTCTGACT	CTCTCAGGGT	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520
	GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GGTGA AAAAT	TGCATCGATG	TAATCAGAAAT	2580
	GAAGACCCGA	TGCCATCCCA	AGATTATGTC	CTCATTATA	ACTATGAGGG	AAGAGGATCT	2640
20	CCAGCTGGTT	TTGTGGGCTG	CTGCAGTGAA	AAGCAGGAAG	AAGATGGCTT	TGACTTTTAA	2700
	AATAATTGGA	AACCCAAAT	TATTACATTA	GCAGAAAGCAT	GCACAAAGAG	ATAATGTACAC	2760
	AGTGCTACAA	TCTGTCTTTT	GTGAGACATT	CTGGAGGTTT	CCAAAAATAA	TATTGTAAAG	2820
	TTCAATTTC	ACATGTATGT	ATATGATGAT	TTTTTCTCA	ATTTTGAATT	ATGCTACTCA	2880
	CCAAATTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAAATTGTA	2940
25	AACAGACAACT	TGGTAAATCT	CAAACTCCAG	CACCTGGAAT	AAGGTCTCTA	AAGCATCTGC	3000
	TCCTTTTTTT	TTTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TGCTCTAATA	TCACATTATT	ATGTATTTCAC	TTTAAAGTAT	AGTTTAAAAA	3120
	ATAACAAGA	AATATTGAGT	ATCACTATGT	GAAGAAAGTT	TTGGAAAGA	AACATGAAG	3180
	ACTGAATTAA	ATTAAAAATG	TTGCAGCTCA	TAAAGAATTG	GGACTCACCC	CTACTGCACT	3240
30	ACCAATTTC	TTTGACTTTG	GAGGC AAAAT	GTGTTGAAGT	GGCCTATGAA	GTAGCAATTT	3300
	TCCTATAGAA	TATAGTTGGA	AATAAATGTG	TGTGTGTATA	TTATTATTAA	TCATGTCAAT	3360
	ATTAAAAATG	AATGAGAGAA	AAAGAGGAAA	ATGGTAAAAA	CTTGAAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGCTGC	3480
	ATTATAACTG	AGTCTATGAG	GAATATGTTT	CTGTCCRAAT	TGTGTAATTT	GTTTAAAAAT	3540
35	GTAAATAAAT	TAAACTTTTC	TGGTTCTGT	GGGAAGGAAA	TAGGGAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCCTGATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACTGCG	TGACAGTGGG	GTTCCCTGCT	TTTTGGTAGC	AAGGCTCCAG	AGATGAGGTG	3720
	TTTTTTTGGG	GGAGCTAATA	ACAAAAACAT	TTTAAACCTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTTTCTATTG	TCTCTTATAG	TGACCAACAT	CTTTTAAAT	TAGATCCAAA	3840
40	TAACCATGTC	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCTCGGG	GAGATTGATT	GTCTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTC	CTCACTGCC	TTCTCTGAG	TGGCAITGGC	4020
	CTGAATCAAG	GAAGCCAGG	CCTTGTGGGC	CCCTTCTTT	CGGCTTCTG	CTAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTTC	ACCATCCTTC	AGCGTGAATT	4140
45	AATTTTTAAT	CAGTTTGCTT	TCTCCAGAGA	AATTTTAAAA	TAATAGAAGA	AATAGA AAT	4200
	TTGAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGAA	CTTTGGGAGA	4260
	AAGCAGCCCA	TAGTGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGCAGAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTTCCTA	GGCTTGGCAC	TGCCTTTTC	TTTCTAGGC	CAATGGCAAC	TGCCATTGTA	4440
50	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTGGGA	GTGAGAAATC	4500
	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACCTG	4560
	GTGTGTGAGA	ACAAACAAAG	CATTATGGG	AATTGTTGTA	TTCTTCTGCG	AGCCCTCCTT	4620
	CTGGGCATTA	AGAGGCTCTA	TGAATTAAT	GCCTATCTAA	AATCTGATT	TATCTCTACA	4680
	TTTTCTGTTT	TCATAATTTA	CCCTAAAATC	TATGTGTTTT	AGACTTAGAC	TTTTTATTGC	4740
55	CCCCCCCCC	TTTTTTTTTG	AGACGGAGTC	TGCTCTGAC	GCACAGGCTG	GAGTGACGTG	4800
	GCTCCGATCT	CTGCTCACTG	AAAGCTCCGC	CTCCCGGGTT	CATGCCATTTC	TCCTGCCCTCA	4860
	GCTCCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCAGC	CCCGCTTAAT	TTTTTGTATT	4920
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTGATCTCC	TGACCTCGTG	4980
	ATCGGCTGCT	CTCGGCTTCC	CAAGGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
	CTGTGTTTCC	GTTTAAAGTC	GTCTTCTTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAGT	5100
60	TGATCATACG	AATTGGATCA	ATCTTGA AAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAAAT	GGAATTCTCT	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCCGTCTGT	AACCAGAGGC	CAGTTTATC	TAAAGGCTAC	TGAAACACCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACCTACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
	CTAGTGCCGA	TAAACTTTCT	CAAAGAGCAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
65	TAACCATCTC	TTTGTCTTTT	GAACATGCTG	AAAACCACT	GGTCTGATG	TATGCCGAAA	5460
	TTTGTAAATC	TTTTCTCTCA	AATGAAAAT	TAAATTTAGG	GATTCAATTC	TATATTTTCA	5520
	CATATGTAGT	ATTATATATT	CCTTATATGT	GTAAGGTGAA	ATTATGGTGA	TTTGAGTGTG	5580
	CAAGAAAATA	TATTTTAAAT	GCTTTCATTT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
70	TGTAAATATA	CAGAATGTTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAATGTCAGT	5700
	GGGTTTGT	TTGCAATGTT	TTAAACAGAG	TTTTAGTATT	GCTATTAAAA	GAAGTTACTT	5760
	TGCTTTTAAA	GAAACTTGCC	TGCTTAA AAT	AAGCAAAAAT	TGGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAAT	ATTAACTTGG	TTTCTGTGTT	TTGCTGTATT	5880
	TAGAGATTAA	ATAATTCTAT	GATGATCACT	TTGCAAAAT	ATGCTTATGG	CTGGCATGGA	5940
75	AATAGAAATA	CTCAATTATG	TCTTTGTGTT	ATTAAATGGG	AATATTTTGG	ACAATGTTTC	6000
	ATTATCAAAAT	TGTGACATC	ATTAAATAT	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT	6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAATT	TTGATCGGGT	6120
	ATTAAAGATA	TTAGAAGGTG	GTATAATTG	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180
	AGGGGTTTTA	CTTTGAGGAC	CAGTGTAGTC	AAGGGAAAAC	ATGAGTTAAA	AAGAAAAGCA	6240
80	GGCAATATTG	CAGTCTTGAT	TCTGCCACTT	ACAGGATAGA	TAATGCCCTG	ACTTTAATGA	6300
	CAGATGATC	CAACCATAGA	GGTGCTCTGT	GCTTCACAGT	GAATCTTTTC	CCCATGCAGG	6360
	AGTGTGCTCC	CCTACAAACG	TTAAGACTGA	TCATTTCAAA	AATCTATTAG	CTATATCAAA	6420
	AGCCTTACAT	TTTAATATAG	GTTGAACCAA	AATTTCAATT	CCAGTAACTT	CTATTGTAAC	6480
	CATTATTTTT	GTTATGTGTT	TCAAGAAATG	TCATTTGATT	TTTGTTTGTA	ATAGTAA AAT	6540
	ACCGGATACA	TTTCACTGTT	CCTTCAGTAT	TGATTTGGTT	GAATATTTGG	TCATAATGGT	6600
85	TGAGAGCAT	GGACACTAGA	GCCAGAAATG	TTGGATATGA	ATCCTGGATC	TGTCATTAC	6660
	TTCTGTGTGA	CCTTTGAAAG	GCTACTTATT	TCCTCTCTTA	GCTTCTCAT	TAAATCAAT	6720
	GAACAATGCC	AGCCTCATGG	GTTGTTGAA	TGATTA AAT	AGTTAATATA	CCTAAGTAC	6780



ATAGAACACT GCCTGCACAT AGTAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840  
 GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTT AGGGAATAA AGTTTGTGCA 6900  
 TATATATAAT CCCGAACAT G

Seq ID NO: 32 Protein sequence:  
 Protein Accession #: NP\_001932.1

1	11	21	31	41	51	
MAAAGPRRSV	RGAVCLHLL	TLVIFSRDGE	ACKKVILNVP	SKLEADKIIG	RVNLEECPRS	60
ADLIRSSDPD	FRVLMDGSVY	TARAVALSDE	KRSFTINLSD	KRKQTQKEVT	VLLHQKQVVS	120
KTRHTRRVL	RAKRRWAPI	PCSMQENSLG	PFPLFLQQVE	SDAAQNYTVF	YSISGRGVDK	180
EPLNLFYIER	DTGNLFCRTP	VDRREYDVFD	LIAYASTADG	YSADLPLPLP	IRVEDENDNH	240
PVFTAIYNF	EVLESSRPGT	TGVVVCATDR	DEPDMHTRL	KYSILQQTTP	SPGLFSVHPS	300
TGVITTVSHY	LDREVVDKYS	LIMKVQDMDG	QFFGLIGTST	CIITVTDSDN	NAPTFRQNAV	360
EAFVEENAFN	VEILRIPIED	KDLINTANWR	VNFTILKQNE	NGHFKISTDK	ETNEGVLSSV	420
KPLNYEENRQ	VNLEIGVNNE	APFARDIPRV	TALNRLVTV	HVRDLDEGPE	CTPAAQYVRI	480
KENLAVGSKI	NGYKAYDPEN	RNGNGLRYKK	LHDPKGMWIT	DEISGSIITS	KILDRREVETP	540
KNELYNITVL	AIDKDDRSCT	GTAVNIEDV	NDNPPEILQE	YVVICPKPMG	YTDILAVDPD	600
EPVHGAPPYP	SLPNTSPBIS	RLWSLTKVND	TAARLSVQKN	AGPQETIPI	TVKDRAGQAA	660
TKLLRVNLCE	CTHPTQCRAT	SRSTGVILGK	WAILAILLGI	ALLFSVLLTL	VCGVFGATKG	720
KRFPEDLAQQ	NLIISNTAP	GDDRVCSSANG	FMTQTNNSS	QGFCTGTMSG	MKNQGQETIE	780
MMKGGNQLE	SCRAGHHHT	LDSCRGGHTE	VDCRYTYSE	WHSPTQPRLG	EKLHRCNQNE	840
DRMPQDVL	TNYNDRGSP	AGSVGCCSEK	QEEDGLDFLN	NLEPKFITLA	EACTKR	

Seq ID NO: 33 DNA sequence  
 Nucleic Acid Accession #: Bos sequence  
 Coding sequence: 64-2583

1	11	21	31	41	51	
GGCAGTCTC	GCTCTCGCA	CCCTCCCGGC	GCCCAGCTTC	TCCTGGCCCT	GCCCAGCATC	60
CGATGGCCG	CCGCTGGCC	CCGGGCTCC	GTGCGCGAG	CCGCTGCTC	GCATCTGCTG	120
CTGACCTCG	TGATCTCAG	TGCTGATGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAATGTA	180
CCTTCTAAC	TAGAGCGAGA	CAAAATAAT	GGCAGAGTTA	ATTTGGAAGA	GTGCTTCAGG	240
TCTGCAGACC	TGATCCGGTC	AACTGATCCT	GATTTACAGG	TTCTAAATGA	TGGCTCAGTG	300
TACACAGCCA	GGGCTGTGTC	GCTGTCTGAT	AAGAAAGAT	CATTTACCAT	ATGGCTTTCT	360
GACAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
TGAGAGACAA	GACACACTAG	AGAACTGTT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCACCT	480
ATTCCTTGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTC	CATTGTTTCT	TCAACAAGTT	540
GAATCTGATG	CAGCACAGAA	CTATCTGTC	TTCTACTCAA	TAAGTGAACG	TGGAGTTGAT	600
AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAATCTATT	TTGCACTCGG	660
CCTGTGGATC	GTGAAGAATA	TGATGTTTT	GATTTGATG	CTTATGCTGC	AACTGCAGAT	720
GGATATTCTG	CAGATCTGCC	CCTCCCACTA	CCCATCAGGG	TAGAGGATGA	AAATGACAAC	780
CACCTGTGTT	TACAGAGAAG	AATTTATAAT	TTTGAAGTTT	TGGAAGTAG	TAGACTGGT	840
ACTACAGTGG	GGTGTGTTG	TGCCACAGAC	AGAGATGAAC	CGACACAAT	GCATACGCGC	900
TGAAATACA	GCATTTTGCA	GCAGACACCA	AGGTCACTCG	GGCTCTTTTC	TGTGCATCCC	960
AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
TCATTGATAA	TGAAGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
ACTTGTATCA	CAGAGTAAC	AGATTCAAAT	GATAATGCAC	CCACTTTTCA	ACAAAATGCT	1140
TATGAAGCAT	TTGTAGAGGA	AAATGCATTC	AATGTGAAA	TCTTACGAAT	ACCTATAGAA	1200
GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCATTTT	AAAGGGAAT	1260
GAAAATGGAC	ATTTCAAAAT	CAGCACAGAC	AAAGAACTA	ATGAAGTGT	TCCTTCTGTT	1320
GTAAAGCCAC	TGAATATGTA	AGAAAACCGT	CAAGTGAACC	TGGAATTTGG	AGTAAACAAAT	1380
GAAGCGCCAT	TGCTAGAGCA	TATTCOCAGA	GTGACAGCCT	TGAACAGAGC	CTTGTTTACA	1440
GTTCATGTGA	GGGATCTGGA	TGAGGGGCTC	GAATGCATCT	CTGCAGCCCA	ATATGTGCGG	1500
ATTAAGAGAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTCATG	ATCTTAAGG	TTGGATCACC	1620
ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GGTTGAAACT	1680
CCCAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAATGATA	ATCCACCAGA	AATACTTCAA	1800
GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAGTTTGC	CCATACTTTC	TCCAGAAATC	1920
AGTAGACTGT	GGAGCCTCAC	CAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
AATGCTGGAT	TTCAAGATA	TACCATTCTT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
GCAACAAAT	TATGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAATCA	GTGCTGTGGG	2100
ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
ATAGCACTGC	TCITTTCTGT	ATTGCTAAT	TTAGTATGTG	GAGTTTITGG	TGCAACTAAA	2220
GGGAACGTT	TTCTGAGAA	TTTAGCACAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTATGA	CCCAACTAC	CAACAACTCT	2340
AGCCAAGGTT	TTTGTGGTAC	TATGGGATCA	GGAATGAAA	ATGGAGGGCA	GGAAACCAT	2400
GAAATGATGA	AAGGAGGAAA	CCAGACCTTG	GAATCTGCTC	GGGGGGCTGG	GCATCATCAT	2460
ACCTGGAAT	CCTGCAGGGG	AGGACACAGC	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520
GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GSTGAAGAAT	CCATTAGAGG	ACACACTGGT	2580
TAAAATTTAA	ACATAAAAGA	AATTCATG	ATGTAATCAG	AATGAAGACC	GCATGCCATC	2640
CCAAGATTAT	GTCTCACTT	ATAACTATGA	GGGAAGAGGA	TCTCCAGCTG	GTTCGTGGG	2700
CTGCTGCACT	GAAAGCAGG	AAGAAGATGG	CCTTGACTTT	TAAATTAATT	TGGAACCCAA	2760
ATTTATTACA	TAGCAGAGG	CATGCACAAA	GAGATAATGT	CACAGTGCTA	CAATTAGGTC	2820
TTTGTGAGAC	ATTCGAGGG	TTTCCAAAAA	TAATATTGTA	AAGTTCAAAT	TCAACATGTA	2880
TGTATATGAT	GATTTTTTTC	TCAATTTTGA	ATTATGCTAC	TCACCAATTT	ATATTTTAA	2940
AGCCAGTTGT	TGCTATCTT	TTCCAAAAAG	TGAAAATGT	TAAACAGAC	AACTGGTAAA	3000
TCTCAAACTC	CAGCACTGGA	ATTAAGGTCT	CTAAAGCATC	TGCTCTTTT	TTTTTTTACG	3060
GATATTTTGA	TAAATTAAT	GCTGGATAAA	TATTAGTCCA	ACAAAGCTA	AGTTATGCTA	3120
ATATCAACTT	ATTATGATT	CACTTAAGT	GATAGTTTAA	AAAATAAACA	AGAAATATTG	3180
AGTATCACTA	TGTGAAGAAA	GTTTTGGAAA	AGAAACAATG	AAGACTGAAT	TAAATTAATA	3240
ATGTTGCAGC	TCATAAGAA	TTGGGACTCA	CCCCTACTGC	ACTACCAAT	TCATTTGACT	3300



TTGGAGGCAA AATGTGTTGA AGTGCCCTAT GAAGTAGCAA TTTTCTATAG GAATATAGTT 3360  
 GGAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTAAAT ATGAATAGAG 3420  
 AACAAAGAGG AAAATGTGTA AAACCTTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA 3480  
 GAAAAAAGAG AGAGCTTCCT AGGCTTGGGC TCTTAAATGC TGCATTATAA CTGAGTCTAT 3540  
 GAGGAAATAG TTCTGTGCCA ATTTGTGTAA TTTGTTTAAA ATTTGTAATA AATTAAACTT 3600  
 TCTGTGTTTC TGTGGGAAGG AAATAGGGAA TCCAATGGAA CAGTAGCTTT GCTTTGCAGT 3660  
 CTGTTCAGG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720  
 GGGGTTCCCT GCTTTTGTGT AGCAAGGGTC CAGAGATGAG GTGTTTTTTT CGGGGAGCTA 3780  
 ATACCAAAAA CATTITAAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTTCTA 3840  
 TTCTCTCTTA TAGTGACCAA CATCTTTTAA ATTTAGATCC AAATAACCAT GTCTCCTTAG 3900  
 AGTTTAGAGG CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAAGCACCTT GGGGAGATTG 3960  
 ATTGTCTTAA AACCTAAGCC CCACAACCTT GACACCTGAT CAGGTCTGGG AGCTACAAAA 4020  
 TTTCAATTTT CTCTCTACTG CCTTCTTCTT GAGTGGCATT GGCTGGAATC AAGGAAAGCC 4080  
 AGGCTTGTG GGGCCCTTTC TTTGGGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATTG 4140  
 CCTTAAGTGA CTCCAGTTT TCCACCATCC TTCAGGCTGA ATTAATTTTT AATCAGTTTG 4200  
 CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTTGAATG TATAAAGAA 4260  
 AAAGATCAAG TTGTCAATTT AGAACAGAGG GAACTTTGGG AGAAAGCAGC CCAAGTAGGT 4320  
 TATTTGTACA GTACAGAGGC AACAGGAAGA TGCAGGCCTT CAAGGGCAAG GAGAGGCCAC 4380  
 AAGGAATAG GGTGGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTT CTAGGCTTGG 4440  
 CACTGCGCTT TCCTTTCTCA GGGCAATGGC AACTGCCATT TGAGTCCGGT GAGGGATCAG 4500  
 CCAACCTCTT CTCTATGGCT CACCTTATTT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560  
 TGCATGATGA GTCTGAAGAG ATTTGCAGGA TGAGCCTGAA CTGGTGTGTG AGAACAAACA 4620  
 AGGCATTCAT GGGAAATGTT GTATTCTTCT TGCAGCCCTC TTTCTGGGCA CTAGAAGGT 4680  
 CTATGAATTA AATGCCTATC TAAATTTCTG ATTTATTCTT ACATTTTCTG TTTTCTAATT 4740  
 TGACCTTAAA ATCTATGTGT TTTAGACTTA GACTTTTAT TTGCCCCCCC CCTTTTCTTT 4800  
 TTGAGAGCGA GTCTGCTCTT GACGCACAGG CTGGAGTGCA GTGGCTCOGA TCTCTGCTCA 4860  
 CTGAAGCTC CGCTCCCGG GTTCATGCCA TTCTCTGCTC TCAGCCTCCT GAGTAGCTGG 4920  
 GACTACAGGC GCCCACGACC ACGCCCGGCT AATTTTGTG ATTTTAAATA GAGACGGGT 4980  
 TTCCTGTGT TAGCCAGGAT GGTCTCGATC TCCTGACCTC GTGATCCGCC TGCCTCGGCC 5040  
 TCCCAAGAGT CTGGGATTAC AGGCATGACC CACCGCTCCC GGCCTTGTGT TCCGTTTAAA 5100  
 GTCTGCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTTGA 5160  
 TCAATCTTGA AATACTCAAC CAAAGACAGG TCGAGAGGCC AGGGGGAGAA AGAAGCTCAG 5220  
 GCACAAATA TTGTCTGAGC AATGGAATTC TCTGTAAGCC TAGTTGCTGA AATTTCTCTC 5280  
 TGTAAACCAGA AGCCAGTTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340  
 CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTCCG ATAACTTTC 5400  
 TCAAGAGCA ACCAGTATCA CTTCCCTGTT TATAAAACCT CTAACCATCT CTTTGTCTT 5460  
 TGAACATGCT GAAACACCAC TGGTCTGCAT GTATGCCOGA ATTTGTAAAT CTTTCTCTC 5520  
 AAATGAAAT TTAATTTATG GGAATTCATT CTATATTTTC ACATATGTAG TATTATTATT 5580  
 TCCITATATG TGTAAAGTGA AATTTATGTT ATTTAGTGT GCAAGAAAT ATATTTTAA 5640  
 AGCTTTTCAAT TTTCCCCAGG TGAATGATTT AGAATTTTT ATGTAAATAT ACAGAATGTT 5700  
 TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAATGCGAG TGGGGTTTGT TTTGCAATGT 5760  
 TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAAACTTGG 5820  
 CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAAGTAATAT TTACAGATGT GGGGAGATGT 5880  
 AATAAACAA TATTAACCTG GCTGCTTAAA ATAAGCAAAA ATTTGATGCA TAAAGTAATA 5940  
 TTTACAGATG TGGGGAGATG TAATAAAACA ATATTAACTT GGTTCCTTGT TTTTGTGTA 6000  
 TTTAGAGATT AAATAATTCT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG 6060  
 GAAATAGAAA TACTCAATTA TGCTTTGTG GTATTAAATG GGAATATTTT GGCAATGTT 6120  
 TCAATATCAA ATGTGCGACA TCATTAATAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180  
 TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATATAATA TTTTGTATCG 6240  
 GTATTAAAA TATTAGAAGG TGGTTATAAT TGCAGATAT TCCATGAATA GTACACTGAC 6300  
 ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAC ACATGAGTTA AAAAGAAAAG 6360  
 CAGGCAATAT TCGAGCTCTG ATCTGCGCAC TTACAGGATA GATAATGCCT GAACTTTAAT 6420  
 GACAAGATGA TCCAACCTTA AAGGTGCTCT GTGCTTCACA GTGAATCTTT TCCCATGCA 6480  
 GGAGTGTGCT CCCCTACAAA GGTAAAGACT GATCATTTC AAAAACTATT AGCTATATCA 6540  
 AAAGCCTTAC ATTTTAATAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600  
 ACCATTATTT TTGTGTATGT CTTCAAGAA GTTCATTGGA TTTTGTGTTG TAATAGTAAA 6660  
 ATACCGGATA CATTTCACGT GTCTTTCAGT ATTGATTGGG TTGAATATTG GGTCAATATG 6720  
 GTTGAGAGGC ATGACACTTA GAGCCAGAA GCTTGGATAT GAATCCTGGA TCTGTCACTT 6780  
 ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCTC TAGCTTCTC ATTAATAATCA 6840  
 ATGAACAATG CCAGCCTCAT GGGGTGTTG AATGATTAATA TTAGTTAATA TACCTAAAGT 6900  
 ACATAGAACCA CTGCGCTGCAC ATAGTAAAAG AATTATAAGT GTGAGGTAGT TGGTAAATTT 6960  
 ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAAGTTTGTG 7020  
 CATATATATA ATCCGAAAC ATG

Seq ID NO: 34 Protein sequence:  
 Protein Accession #: NP\_077741.1

1 11 21 31 41 51  
 MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60  
 ADLIRSSDPD FRVLNDGSVV TARAVALSDK KRSFTINLSD KRKQTQKEVT VLLHQQKQVS 120  
 KTRHRTRETVL RRAKRRWAPI PCSMQENSLG PPPLFLQQVE SDAQNTTVF YSISGRVDK 180  
 EPLNLFYIER DTGNLPCTRP VDREYDVDP LIAYASTADG YSADLPLPLP IRVEDENDNH 240  
 PVFTEALYNF EVLESSRPGT TVGVVCATDR DEPDMTETRL KYSILQOTPR SPGLFSVHPS 300  
 TGVITTVSHY LDREVVDKYS LIMKVQMDMG QFFGLIGTST CIITVTDSDN NAPTFRQNAV 360  
 EAFVEENAFN VEILRIPIED KDLINTANWR VNFTILKQNE NGHFKISTDK ETNEGVLSVV 420  
 KPLNYEENRQ VNLEIGVWNE AFFARDIPRV TALNRLVTV HVRDLDEGPE CTPAQYVRI 480  
 KENLAVGSKI NGYKAYDPEN RENGRLRYKK LHDPKGWITT DEISGSIIIS KILDRREVETP 540  
 KNELYNITVL AIDKDRSCT GTLAVNIEDV NDNPPPELQE YVVICPKMG YTDILAVDPD 600  
 EPVHGAPFYF BLPTTSPEIS RLWSLTQVND TAARLSYQKN AGFQYETIPI TVKDRAGQAA 660  
 TKLLRVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720  
 KRFPEDLAQQ NLIISNTEAP GDRVCSANG FMTQTNNSS QGFGTMSGG MKNGGQETIE 780  
 MMKGNGTLE SCRGAHHHT LDSCRGHTE VDNCRYTSE WBSFTQPRLG EESIRGHTG

Seq ID NO: 35 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 146-1273

1 11 21 31 41 51  
5 GGGAGTGGGC GTGGCGGTGC TGCCAGGTG AGCCACCGCT GCTTCTGCC AGACACGGTC 60  
GCCTCCACAT CCAGGTCTTT GTGCTCCTCG CTGCGCTGTT CCTTTTCCAC GCATTTTCCA 120  
GGATAACTGT GACTCCAGGC CCGCAATGGA TGCCCTGCAA CTAGCAAATT CGGCTTTTGC 180  
CGTTGATCTG TTCAAACAAC TATGTGAAAA GGAGCCACTG GGCAATGTCC TCTTCTCTCC 240  
AATCTGTCTC TCCACTCTC TGTCACCTGC TCAAGTGGGT GCTAAAGGTG ACACGTGCAA 300  
10 TGAATTTGGA CAGGTTCTTC ATTTTAAAA TGTCAAAGAT ATACCCTTTG GATTTCAAAC 360  
AGTAACATCG GATGTAAACA AACTTAGTTC CTTTACTCA CTGAACTAA TCAAGCGGCT 420  
CTACGTAGAC AAATCTCTGA ATCTTTCTAC AGAGTTCATC AGCTCTACGA AGAGACCCTA 480  
TGCAAGGAA TTGAAACTCG TTGACTTCAA AGATAAATG GAAGAAACGA AAGGTCAGAT 540  
CAACAACFCA ATTAAGGATC TCACAGATGG CCACTTTGAG AACATTTTAG CTGCAACAG 600  
TGTGAACGAC CAGACCAAAA TCCTTGTGGT TAATGCTGCC TACTTTGTTG GCAAGTGGAT 660  
15 GAGAAATTTT CCGTAATCAG AAACAAAAGA ATGTCCTTTC AGACTCAACA AGACAGACAC 720  
CAACCAAGTG ACATGATGA ACATGGAGGC CACGTTCTGT ATGGGAAACA TTGACAGTAT 780  
CAATTGTAAG ATCATAGAGC TTCCTTTTCA AAATAAGCAT CTCAGCATGT TCATCCTACT 840  
ACCCAGAGAT GTGGAGGATG AGTCCACAGG CTGGGAGAAG ATTGAAAAC AACTCAACTC 900  
20 AGAGTCACTG TCACAGTGGG CTAATCCAG CACCATGGCC AATGCCAAGG TCAAACTCTC 960  
CATTCCAAAT TTTAAGGTGG AAAAGATGAT TGATCCCAAG GCTTGTCTGG AAAATCTAGG 1020  
GCTGAAACAT ATCTTCAGTG AAGACACATC TGATTTCTCT GGAATGTCAG AGACCAAGGG 1080  
AGTGGCCCTA TCAATGTTA TCCACAAAGT GTGCTTAGAA ATAACTGAAG ATGGTGGGGA 1140  
TTCCATAGAG GTGCCAGGAG CACGATCTCT GCAGCACAAG GATGAATTGA ATGCTGACCA 1200  
25 TCCTTTTAT TACATCATCA GGCACAAACA AACTCGAAAC ATCATTTTCT TTGGCAAAAT 1260  
CGTTTCTCT TAAGTGGGAT AGCCCATGTT AAGTCTCTCC TGACTTTTCT GTGGATGCGG 1320  
ATTTCTGTAA ACTCTGCATC CAGAGATTCA TTTTCTAGAT ACAATAAAT GCTAATGTTG 1380  
CTGGATCAGG AAGCCGCCAG TACTTGTCTAT ATGTAGCCTT CACACAGATA GACCTTTTCT 1440  
TTTTTCCAAAT TCTATCTTTT GTTTCCTTTT TTCCCAATAG ACAATGACAT ACGCTTTTAA 1500  
30 TGAAGAAGAA TCAGCTTAGA GGAATAATAT TTATTCAAT TTTGTCAAAT TGTCCGGGGT 1560  
AGTTGGCAGA AATACAGTCT TCCACAAAGA AAATTCCTAT AAGGAAGATT TGAAGCTCT 1620  
TCTTCCAGC ACTATGCTTT CTTCTTTGG GATGAGAAT GTTCCAGACA TTCTCGCTTC 1680  
CCTGAAAGAC TGAAGAAATG GTAGTGCATG GGAACCCAGA AACTGCCCTG GCTCCAGTGA 1740  
AACTTGGGGA CATGCTCAGG CTACTATAGG TCCAGAAGTC CTTATGTTAA GCCCTGGCAG 1800  
35 GCAGGTGTTT ATTAATAATTC TGAATTTTGG GGATTTTCAA AAGATAATAT TTTACATACA 1860  
CTGTATGTTA TAGAACTTCA TGGATCAGAT CTGGGGCAGC AACCTATAAA TCAACACCTT 1920  
AATATGCTGC AACAAATGT AGAATATTCA GACAAATGG ATACATAAAG ACTAAGTAGC 1980  
CCATAAGGGG TCAAAATTG CTGCCAAATG CGTATGCCAC CAACCTACAA AAACACTTCG 2040  
TTCCAGAGC TTTTCAGATT GTGGAATGTT GGATAAGGAA TTATAGACCT CTAGTAGCTG 2100  
40 AAATGCAAGA CCCCAGAGG AAGTTCAGAT CTTAATATAA ATTCACITTC ATTTTGTATA 2160  
GCTGTCCCAT CTGGTCACTG GGTGGGCACT AGACTGGTGG CAGGGGCTTC TAGCTGACTC 2220  
GCACAGGAT TCTCACAATA GCCGATATCA GAATTTGTGT TGAAGGAAT TGTCTCTTCA 2280  
TCTAATATGA TAGCGGGAAG AGGAGAGGAA ACTACTGCCT TTAGAAAATA TAAGTAAAGT 2340  
GATTAAGATG CTCACGTTAC CTTGACACAT AGTTTTCAG TCFATGGGTT TAGTTACTTT 2400  
45 AGATGCAAG CATGTAACCT ATATTAAATG TAATTTGTAA AGTTGGGTGG ATAACTATC 2460  
CCTGTGGCG GTTCATGGAT TACTTCTCTA TAAATAATAT ATATTACCA AAAAATTTTG 2520  
TGACATTCCT TCTCCCATCT CTTCCCTGAC ATGCATTGTA AATAGGTTCT TCTTGTCTG 2580  
AGATTCAATA TTGAATTTCT CCTATGCTAT TGACATAAA ATATTATTGA ACTACC

Seq ID NO: 36 Protein sequence:  
Protein Accession #: NP\_002630.1

1 11 21 31 41 51  
55 MDALQLANSA FAVDLFKQLC EKEPLGNVLF SPICLSTSLA LAQVGAGDGT ANEIGQVLHF 60  
ENKDIIPFPG QTVTSNVKL SSFYSKLKLIK RLYVDKSLNL STEFISSTKR PYAKELETVD 120  
FKDKLEETKG QKINSIKDLT DGHFENILAD NSVNDQTKIL VVNAAYFPVG WMKFPPESET 180  
KECPFRINKT DTFQVMNMN EATFCMGNID SINCKIIELP FQNKHLSMFI LLPKDVESD 240  
TGLEKIEKQL NSESLSQWN PSTMANAKVK LSIPKPKVEK MIDPKACLEN LGLKHIFSED 300  
60 TSDPSGMSSET KGVALSNIH KVCLEITEDG GDSIEVPGAR ILQHKDELNA DHPFIYIIRH 360  
NKTRNIIFPG KFCSP

Seq ID NO: 37 DNA sequence  
Nucleic Acid Accession #: NM\_0168583  
Coding sequence: 72-842

1 11 21 31 41 51  
70 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60  
TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCCA 120  
CCATGGCCCA GTTTGGAGGC CTGCCGCTGC CCGCTGGACCA GACCCCTGCC TTGAATGTGA 180  
ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGAGGAAAG CTGACAAAT GCCCTCAGCA 240  
ATGGCTGCT GTCTGGGGGC CTGTTGGGCA TTCTGGAAAA CCTTCCGCTC CTGGACATCC 300  
75 TGAAGCCTGG AGGAGTACT TCTGTTGGCC TCCTTGGGGG ACTGCTTGA AAAGTGACGT 360  
CAGTGATTCC TGGCCCTGAAC AACATCATG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420  
AACTTGGCCT TGTGAGAGC CCTGATGGCC ACOGTCTCTA TGTCAACATC CCTCTCGGCA 480  
TAAAGTCCA AGTGAATACG CCGCTGGTGG GTGCAAGTCT GTTGGAGCTG GCTGTGAAGC 540  
TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGGTCC 600  
80 TTGGTGACTG CACCCATTCC CCGTGAAGCC TGCAAAATTC TCTGCTTGAT GGAAGTGGCC 660  
CCCTCCCAT TCAAGGTCTT CTGGACAGCC TCACAGGAT CTTGAATAAA GTCTGCTG 720  
AGTTGGTTCA GGGCAACGTG TGCCCTCTGG TCAATGAGT TCTCAGAGGC TTGACATCA 780  
CCCTGGTGCA TGACATTGTT AACATGCTGA TCCACGACT ACAGTTTGTG ATCAAGGTCT 840  
AAGCCTTCCA GGAAGGGGCT GGCCTCTGCT GAGCTGCTTC CAGTGTCTCA CAGATGGCTG 900  
85 GCCATGTGC TGAAGATGA CACAGTGGC TTCTCTCGGA GGAACCTGCC CCTCTCTCTT 960  
TCCACACAGG CGTGTGTAAC ATCCATGTG CTTCACTTAA TAAATGGTCT CTCTCTCTG 1020  
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA

## WO 02/086443

Seq ID NO: 38 Protein sequence:  
Protein Accession #: NP\_057667

PCT/US02/12476

5 1 11 21 31 41 51 60  
MFQTGGILVF YGLLAQTMAQ FGGLFVPLDQ TLPLNVNPL PLSPTGLAGS LTNALSNGLL 120  
SGLLGLILEN LPLLDILKPG GGTSGGLLGG LLGKVTSVIP GLNNIIDIKV TDPQLLELGL 180  
VQSPDGHRLY VTIPLGKILQ VNTPLVGASL LRLAVKLDIT ABILAVRDQ ERIHLVLGDC 180  
THSPGSLQIS LLDGLGPLPI QGLLDSLTGI LANKVLPVLQ GNVCPVNEV LRGLDITLVH 240  
10 DIVNMLIHGL QFVIKV

Seq ID NO: 39 DNA sequence  
Nucleic Acid Accession #: NM\_004363.1  
Coding sequence: 115-2223

15 1 11 21 31 41 51 60  
CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 120  
TCTCGGAAGT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 180  
TCTCCCTCGG CCCCTCCCCA CAGATGGTGC ATCCCTGCGC AGAGGCTCCT GCTCACAGCC 180  
TCACTTCTAA CCTTCTGGAA CCGGCCACC ACTGCCAAGC TCACTATTGA ATCCAACGCG 240  
TTCAATGTGC CAGAGGGGAA GGAGGTGCTT CTACTTGTC ACAATCTGCC CCAGCATCTT 300  
TTTGGCTACA GCTGGTACAA AGGTGAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360  
GTAATAGGAA CTCAACAGC TACCCAGGG CCGGCATACA GTGGTGGAGA GATAATATAC 420  
CCCAATGCAT CCGCTCTGAT CCAGAACATC ATCCAGAAATG ACACAGGATT CTACACCCCTA 480  
CAGCTCATAA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCOG GGTATACCCG 540  
GAGCTGCCCA AGCCCTCCAT CTCCAGCAAC AACTCCAAAC CCGTGGAGGA CAAGGATGCT 600  
GTGGCCTTCA CCTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT 660  
CAGAGCCTCC CGGTCACTCC CAGGCTGCAG CTGTCCAAATG GCAACAGGAC CCTCACTCTA 720  
TTCAATGTCA CAGGAAATGA CACAGCAAGC TACAAATGTG AAAACCCAGAA CCCAGTGAGT 780  
GCCAGGCGCA GTGATTCAGT CATCTGAAT GTCCCTCTATG GCCCGGATGC CCCACCAATT 840  
TCCCTCTTAA ACACATCTTA CAGATCAGGG GAAATCTGA ACCTCTCTG CCACGAGGCC 900  
TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAGCA ATCCACCCAA 960  
GAGCTCTTTA TCCCAACAT CACTGTGAAT AATAGTGGAT CCTATAGTG CCAAGCCCAT 1020  
AACTCAGACA CTGGCCTCAA TAGGACACA GTCAAGCAGA TCACAGTCTA TGCAGAGCCA 1080  
CCCAACCCCT TCATCACCAG CAACAACCTC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140  
TTAAGCTGTG AACCTGAGT ACATCTCTT GCTGATGAT GGGAAACATC AGCAACACAC ACAAGAGCTC 1200  
CTCCCGGTCA GTCCAGGCT GCAGCTGTCC AATGACAACA GGACCTCAC TCTACTCAGT 1260  
GTCAACAGGA ATGATGTAGG ACCCTATGAG TGTGGAATCC AGAAGCAATT AAGTGTGAC 1320  
CACAGGACCC CAGTCTCCTT GAATGTCTCT TATGGCCGAC ACGACCCAC CATTTCGCC 1380  
TCATACACTT ATTACGCTCC AGGGGTGAAC CTCAGCCTCT CCGCCATGTC AGCCTCTAAC 1440  
CCACTGCAC AGTATTCTTG GCTGATGAT GGGAAACATC AGCAACACAC ACAAGAGCTC 1500  
TTTATCTCCA ACATCACTGA GAAGAACAGC GGAATCTATA CCGCCAGGC CAATAACTCA 1560  
GCCAGTGACC ACAGCAGGAC TACAGTCAAG ACAATCACAG TCTCTGGGA GCTGCCCAAG 1620  
CCCTCCATCT CAGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GSCCTTCACC 1680  
TGTGAACCTC AGGCTCAGAA CACAACCTAC CTGTGTGGG TAAATGGTCA GAGCCTCCCA 1740  
GTCACTCCCA GCGTCACTG GTCCAATGGC AACAGGACCC TCACTCTATT CAATGTGACA 1800  
AGAAATGACG CAGAGAGCTA TGTATGTGGA ATCCAGAACT CAGTGTGTC AAACCGCAGT 1860  
50 GACCCAGTCA CCGTGGATGT CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCAGAC 1920  
TGCTCTTACC TTTGGGAGC GAACCTCAAC CTCTCTGACC ACTCGGCTC TAACCCATCC 1980  
CCGAGTATT CTGGCGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTTATC 2040  
GCCAAATCA CGCCAAATA TAACGGGACC TATGCTCTGT TTGCTCTTAA CTGGCTACT 2100  
GGCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TCTCTCTGGT 2160  
55 CTCTCAGCTG GGGCCACTGT CGGCATCATG ATTGAGTGC TGGTTGGGT TGCTCTGATA 2220  
TAGCAGCCCT GGTGTAGTCT CTTCATTCA GGAAGACTGA CAGTTGTTTT GCTTCTTCT 2280  
TAAGCAATTT GCAACAGCTA CAGTCTAAAA TTGCTTCTT ACCAAGGATA TTTACAGAAA 2340  
AGACTCTGAC CAGAGATCGA GACCATCCTA GCCAACATCG TGAACCCCA TCTCTACTAA 2400  
60 AAATCAAAA ATGAGCTGGG CTGGTGGG OGCACTCTGA GTCCAGTTA CTCGGGAGGC 2460  
TGAGGCAGGA GAATGCTTG AACCCGGGAG GTGGAGATTG CAGTGAGCCC AGATCGCACC 2520  
ACTGCACTCC AGTCTGGCAA CAGAGCAAGA CTCATCTCA AAAAGAAAAG AAAAGAAGAC 2580  
TCTGACCTGT ACTCTGAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640  
AACTTTAATG AACTAACTGA CAGCTTCATG AAACGTGTCCA CCAAGATCAA GCAGAGAAAA 2700  
TAATTAATTT CATGGGACTA AATGAATCAA TGAGGATGTC TGATTTCTTA AATGTCTTGT 2760  
65 TTCCAGATT TTGTGAACAA AAATTGAGAC ATTTACATTT TCTCCCTATG TGGTCGCTCC 2880  
AAATATACCT TTGTGAACAA AAATTGAGAC ATTTACATTT TCTCCCTATG TGGTCGCTCC 2880  
AGACTTGGGA AACTATTAT GAATATTAT ATTGTATGTT AATATAGTTA TTGCACAAGT 2940  
TCAATAAAAA TCTGCTCTTT GTATAACAGA AAAA

Seq ID NO: 40 Protein sequence:  
Protein Accession #: NP\_004354.1

75 1 11 21 31 41 51 60  
MESPSAPPHR WCIPWQRLLL TASLLTFWNP PITAKLTIES TPFNVAEKGE VLLLVHNLPO 60  
HLFGYSWKYG BRVDGNRQII GYVIGTQQAT PGPAYSGREI IYFNASLLIQ NIIQNDTGFY 120  
TLHVIKSDLV NEBATGQFRV YPELPKPSIS SNNSKPVEDK DAVAPTCEPE TQDATYLVWV 180  
NNQSLFVSPR LQLSNGNRTL TLFNVTRNDT ASYKCEIQNP VSARRSDSVI LNVLYGPDAP 240  
80 TISPLNTSYR SGENLNLSCH AASNPPAQYS WFNVTGQQS TQELFIPNIT VNNSGSYTCQ 300  
AHNSDTGLNR TTVTTIVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ NTTYLWNVNN 360  
QSLFVSPRLQ LSNDRNLTLT LSVTRNDVGP YECGIONELS VDHSDPVILN VLYGDDPTI 420  
SPSYTYRPG VNLSLSCHAA SNPPAQYSWL IDGNIQQHTQ ELFISNITEK NSGLVTCQAN 480  
NSASGHSRTT VKTITVSAEL PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWVWNGQS 540  
LPVSPRLQLS NGNRTLFLPN VTRNDARAYV CQIQNSVSAN RSDPVTLDLV YGPDPTIISP 600  
85 PDSLYLGGAN LNLSCSASN PBPQYSWRIN GIPQOHTQVL FIAKITPNNN GTYACFVSNL 660  
ATGRNNSIVK SITVSASGTS PGLSAGATVG IMIGVLVGVA LI

Seq ID NO: 41 DNA sequence  
Nucleic Acid Accession #: NM\_006952.1  
Coding sequence: 11-793

5  
10  
15  
20  
1 11 21 31 41 51  
AATCCGACAA ATGGCGAAAG ACAACTCAAC TGTTCGTTGC TTCCAGGGCC TGCTGATTTT 60  
TGGAATGTG ATTATTGGTT GTTGCAGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120  
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180  
GGCTGCCTGG ATCGGCATAT TTGTGGGCAT CTGCCTCTTC TGCCCTGCTG TCTAGGSCAT 240  
TGTAAGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTTCATT TGATGTTTAT 300  
AGTATATGCC TTGAAGTGG CATCTGTAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360  
ACCCAACCTC TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAACGCC CTCCAAACAA 420  
TGATGACCAG TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480  
CAATTGCTGT GGGCTAAATG GTCCATCAGA CTGGCAAAA TACACATCTG CCTTCCGGAC 540  
TGAGAAATAA GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600  
AGAACCTCTC AACCTGGAGG CTGTAAACT AGGCGTGCTT GGTTTTATC ACAATCAGGG 660  
CTGCTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CCTGTTTGG 720  
ATTGTCATT CTCTGCTGGA CTTTTTGGGT TCTCTGGGT ACCATGTTCT ACTGGAGCAG 780  
AATTGAATAT TAAGAA

Seq ID NO: 42 Protein sequence:  
Protein Accession #: NP\_008883.1

25  
30  
1 11 21 31 41 51  
MAKDNSTVRC FQGLLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAM 60  
IGTFVGLCLF CLSVLGIYGI MKSSRKILLA YPILMFIVYA FEVASCITAA TQRDFFTPNL 120  
FLKQMLERYQ NNSPNNDDQ WKNGVTKTW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180  
DADYFWPRQC CVMNKLKEPL NLEACKLGVP GFYHNQGCYE LISGPMNRHA WGVAWFGFPI 240  
LCWTFWVLLG TMFYWSRIEY

Seq ID NO: 43 DNA sequence  
Nucleic Acid Accession #: E05 sequence  
Coding sequence: 83-2605

40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
1 11 21 31 41 51  
GCCGACAGAA TCTGCGCGTA TCCTGGAGCC GGCCAGTTG TGAAGTAGGA GAGCTTTGGG 60  
ACCTCTGTCC CAAGCAAGAG AGATGAATGG AGAGTATAGA GGCAGAGGAT TTGGACGAGG 120  
AAGATTTCAA AGCTGGAAGA GGGGAGAGGG TGGTGGGAAC TTCTCAGGAA AATGGAGAGA 180  
AAGAGAACAC AGACCTGATC TGAGTAAAC CACAGGAAAA CGTACTTCTG AACAAACCCC 240  
ACAGTTTTTG CTTTCAACAA AGACCCACCA GTCAATGCAG TCAACATTGG ATCGATTTCAT 300  
ACCATATAAA GGCCTGGAGC TTTATTTCTC TGAAGTTTAC AGCGATAGCT CTCTTTTGAT 360  
TGAGAAGATT CAAGCATTTG AAAAATTTT CACAAGGCAT ATTGATTGTG ATGACAAGGA 420  
TGAAATAGAA AGAAAGGGAA GTATTTTGGT AGATTTTAAA GAACTGACAG AAGGTGGTGA 480  
AGTAACCTAC TTGATACCAAG ATATAGCAAC TGAACCTAAGA GATGCACCTG AGAAACCTT 540  
GGCTTGCAAT GGTTTGGCAA TACATCAGGT GTTAACTAAG GACCTTGAAA GGCAATGCAG 600  
TGAGTTACAA GGCCTGGAGG GATTGTCTAA TGATGGAGAA ACAATGGTAA ATGTGCCACA 660  
TATTCATGCA AGCGGTGACA ACTATGAGCC TTTGACACAG CTCAGAATG TCAGAGCAAA 720  
TTACTATGGA AAATACATTG CTCTAAGAGG GACAGTGGTT CGTGTACGTA ATATAAGCC 780  
TCTTTGCACC AAGATGGCTT TCTTTGTGTC TGCATGTGGA GAAATTCAGA GCTTTCTCT 840  
TCCAGATGGA AAATACAGTC TTCCCAACAA GTGTCTCTGT CCGTGTGTGTC GAGGCAGGTC 900  
ATTTACTGCT CTCGCGAGCT CTCTCTCAC AGTTACGATG GACTGGCAGT CAATCAAAAT 960  
CCAGGAATTG ATGTCTGATG ATCAGAGAGA AGCAGGTGCG ATTCCACGAA CAATAGAATG 1020  
TGAGCTTGTT CATGATCTTG TGGATAGCTG TGTCCCGGGA GACACAGTGA CTATTACTGG 1080  
AATTGTCAAA GTCTCAATG CGGAAGAAGG TTCTCGAAT AAGAATGACA AGTGTATGTT 1140  
CCTTTTGTAT ATTGAAGCAA ATTTCTATTG TAATAGCAAA GGACAGAAAA CAAAGAGTTC 1200  
TGAGGATGGG TGTAAGCATG GAATGTTGAT GGAGTTCTCA CTTAAAGACC TTTATGCCAT 1260  
CCAAGAGATT CAAGCTGAAG AAAACCTGTT TAAACTCATT GTCAACTGCG TTTGCCCTGT 1320  
CATTTTGTGT CATGAACCTG TTAAGCAGG TTTGGCATTA GCACCTCTTG GAGGAAGCCA 1380  
GAATACGCGA GATGACAAA ACAGAATTCC AATTGGGGGA GACCCCAACA TCCTTGTGTT 1440  
TGGAGATCCA GGCCTAGGAA AAGTCAAAT GCTACAGGCA GCGTGCATG TTGCCCCACG 1500  
TGGCGTGTAT GTTTGTGGA ACACCCAGAC CACCTCTGCT CTGACGGTAA CTCTTTCAAA 1560  
AGATAGTTCC TCTGGAGATT TTGCTTTGGA AGCTGGTGCC CTGTTACTTG GTGATCAAGG 1620  
TATTTGTGGA ATCGATGAAT TTGATAAGAT GGGGAATCAA CATCAAGCCT TGTGGAAGC 1680  
CATGGAGCAG CAAAGTATTA GTCTTGCTAA GGCTGGTGTG GTTTGTAGCC TTCTGCAAG 1740  
AACTTCCATT ATTGCTGCTG CAAATCCAGT TGGAGGACAT TACAATAAAG CCAAAACAGT 1800  
TTCTGAGAA TTAATAATGG GGAGTGCACT ACTATCCAGA TTTGATTGCG TCTTTATCCT 1860  
GTTAGATACT CCAATAGAGC ATCATGATCA CTACTCTCT GAACATGTGA TTGCAATAAG 1920  
AGCTGGAAG CAGAGAACCA TTAGCAGTGC CACAGTAGCT CGTATGAATA GTCAAGATT 1980  
AAATCTTCC GTACTTGAAG TAGTTTCTGA GAAGCCATTA TCAGAAAGAC TAAAGGTGTT 2040  
TCTGTGAGAA ACAATAGATC CCATTCCCA CAGCTATTG AGAAAGTACA TTGGCTATGC 2100  
TGGCAGTAT GTGATCCCAA GGCTATCCAC AGAAGCTGCT CGAGTCTTTC AAGATTTTAA 2160  
CCTTGAGCTC CGGAAACAGA GCCAGAGGTT AAATAGCTCA CCAATCACTA CCAAGCAGCT 2220  
GGAATCTTTG ATTCGTCTGA CAGAGGCAGC AGCAAGGTTG GAATTGAGAG AGGAAGCAAC 2280  
CAAAGAAGAC GCTGAGGATA TAGTGGAAAT TATGAAATAT AGCATGCTAG GAATCTACT 2340  
TGATGAATT GGGAACTTAG ATTTTGAAG ATCCAGCAT GGTTCGGAAG TGAGCAACAG 2400  
GTCAACAGCG AAAAGATTAA TTTCTGCTCT CAACAACGTT GCTGAAGAA CTTATAATAA 2460  
TATATTTCAA TTTCTATCAAC TTCCGAGAT TGCCAAAGAA CTAAACATTC AGGTGTCTGA 2520  
TTTTGAAAT TTTATGGGAT CACTAAATGA CCAGGGTTAC CTCTGAAAA AAGGCCCAAA 2580  
AGTTTACCAG CTTCAACTA TGTAAAAGGA CTTCACCAAG TTAGGGCCTC CTGGGTTTAT 2640  
TGCAGATTAA AGCCATCTCA GTGAAGATAT GCGTGCAACG ACAGACAGAC AGACACACAC 2700  
ACACACACAC ACACACACAC ACACACACAC ACACACAGTC AAATACTGTT CTCTGAAAAA 2760  
TGATGTCCCA AAGTATTAT AATAGGAAAA AAGCATTAAA TATAATAAAC TAATTTAAGA 2820

AGTGATAAAG TCCTCAGATG CAGTAGCTCA CACTGTAATC ACAGTGACTC AGGAGGCTGA 2880  
GGTGAGAGGA TTCCTTGAGG CCAGGGTTCG AGACCAACCT TGGGCAACAT AGCAGAGCCC 2940  
CATTCTCTAA AAAAAAAAAA AAAAAATTTA AACTTAGCTG GGTATGGTGG CACATGCCTA 3000  
TAGCTCTAGC TACTGTGAGG GCTGAGGCAG GAGGATTCTT TGAGCCAGG AGTTTGAGGT 3060  
TACAGTGAGC CACAATCACA CCAATCACTG CACTCCAGCC TGGGCAATAA AGTAACTCTT 3120  
GACTCAAAA AATAAAAAAA ATTGTAGTGG TAGCCATGTG TTAATTGTTA AATAAATTCT 3180  
CCAAAGGGCT AAAAGTAAAT TACTTATAAA TTTTATTATG TTGTATTTTT GACCTGCCTT 3240  
TTATATGTAT GAATATTCCA TAGTTTTGCA TATCAGATGT AGGCATACAG ACAATATCAT 3300  
AAACCAATGA ATATATTACA TATTCTGTGT TCCAATAAAA CTTTATTATG GGACACTAAA 3360  
ATTGGAATTT CATAAAATTT TCCCATGTCA AGAATACAAA ATACTTGAGT TTTGTTTTTA 3420  
GCTATTTAAT AATAGGCTCTC ATTATTCCCA CAGGCTGTAG TTTGTAGTCT TGCTTGAAC 3480  
AATAGAAACA GACTGATTAA GCAGGAGAAG TTTTGTGAAA GAATTTGTGT TGGCTCACGG 3540  
AATTATTAGA AGGCAGGTGA ACCAGGAGGG TAAGCTTCCA GCAGCAATTT GTAAAAACAT 3600  
GCCTTAGAAT TGGACTAAGG AAGAAGCTGC TGACACTCCA CTGCCACACA GGGCACTGGA 3660  
AGAAAGTGCT GCTGCTCTCC TGCCCACTCT TGCCACAGGA ATAGGTAGAA 3720  
GAATGCCCCC ACCCGACCCG GAACAGCAAC AAAAGGATTC TGCATGAGAT GCCTCCCTAA 3780  
ATTGCTGAAT TCAAAAAGA AGTTGCATAC AAAGACATCT GATTGAAAAA GGGTATGTTA 3840  
TAGTCCCTTT TCATAGGCTG CTAGGGAGTT TTCCTGGTTC TACTTTCAGG TGGTGGGATC 3900  
AATAAGACCA GAATTTCTCA TATGTTGTGA GAGGATTCAG ATGTTACAGG GTTGCCAGCC 3960  
AACTATCAA TCATGTATAA ATCCAACAAA CACTTTGTAA CATACAAGAA CTCAGGAAAT 4020  
GTGAACCAT GTTGAGATTA CTACTAAAAAT ACGGCTTCCC GCAAAACGAAG ATGAATGGAA 4080  
AATGTAAATA AAAAGAACTG GCAGTGTATA TCAGATGTTT AACTATAGGA CCAGAACTAA 4140  
GATGTGGAGA CTATTGCCAT AGACCACAAT GTAAATTTTT AAGTGAGGAA GGAATAATCA 4200  
GGAATCAAAA GGGGCCAGGT GCAGTGGCTC ACATCTATAA TCCCAGAGCT TTGGGAGTTC 4260  
GAGGCAGAGG GATCACTTGA AGCCAGTTTT GAGACCAAGC TATGCAACAC ATTGAGACCC 4320  
TATCTCTACA AAAAAATAGT TAGCTGGGCA CGGTGGTGCA TGCCTATTGT CCTACCTACT 4380  
GTGGAGGCTG AAGTAGGAAA TCACTTGAGC CCGAGAGTTT GAGGTTACAG TGAGCTATGA 4440  
TTATACCACT GCACTCCAGC CTGGGCAAGA GAGCAAGACC TTGTCTCTT

Seq ID NO: 44 Protein sequence:  
Protein Accession #: CAB55276.2

1 11 21 31 41 51  
MNGEYRGRGP GRGRFQSWKR GRGGGNFSGK WREREHRPDL SKTTGKRTSE QTPQFLLSTK 60  
TPQSMQSTLD RPPIPKGWKL YPSEVYSDSS PLIEKIQAPE KPFTRHIDLY DKDEIERKGS 120  
ILVDFKELTE GGEVNLNLPD IATELRDAPE KTLACMGLAI HQVLTKDLER HAAELQAQEG 180  
LSNDGETMVN VPHIGARVYN YEPLTQLKNV RANYYGKYLIA LRGTVVVRVSN IKPLCTKMAF 240  
LCAACGBIQS FPLPDGKYSL PTKCPVPVCR GRSFTALRSS PLTVTMDWQS IRIQELMSDD 300  
QREAGRIPRT IECELVDELV DSCVPGDTVT ITGIVKVSNA EDSRNKNNDK CHFLLYIEAN 360  
SISNSKGQKT KSEEDGCKHG MLMEFSLKDL YAIQEIQAEE NLPLKLIIVNSL CPVIFGHELV 420  
KAGLALALFT GSQKYADDKN RIPIRGDPHI LVVDPDPLGK SQMLQAACNV AFRGVVYVCGN 480  
TTTTSGLTVT LSKDSSSGDF ALEAGALVLG DQGIQIDEF DKMGNHQHAL LEAMEQQSIS 540  
LAKAGVVCSL PARTSLIAAA NPVGGHYNKA KTVSENLMKG SALLSRPDLV FILLDTPNEH 600  
HDELLSEHVI AIRAGKQRTI SSATVARMNS QDSNTSVLEV VSEKPLSERL KVVPGETIDP 660  
IPHQLLRXYI GYARYVYVPR LSTEARVLQ DFYLELRKQS QRLNSSPITT RQLESILRLT 720  
BARARLEIRE EATKEDAEDI VEIMKYSMLG TYSDEFGNLD FERSQEGSGM SNRSTAKRPI 780  
SALNNVAERT YNNIFQFEQL RQIAKELNIQ VADFENFIGS LNDQGYLLKK GKRVYQLQTM

Seq ID NO: 45 DNA sequence  
Nucleic Acid Accession #: NM\_005416.1  
Coding sequence: 149..658

1 11 21 31 41 51  
ACCAGATCCC AGAGGCTGAA CACCTCGACC TTCTCTGCAC AGCAGATGAT CCCTGAGCAG 60  
CTGAAGACCA GAAAGGCCAC TAAGACTTTC TGCTTAATTC AGGAGCTTAG AGGATTCTTC 120  
AAAGAGTGTG TCCACGATCC TTTGAAGCAT GAGTTCTTAC CAGCAGAAGC AGACCTTTAC 180  
CCCACCACCT CAGCTTCAAC AGCAGCAGGT GAAACAACCC AGCCAGCCTC CACCTCAGGA 240  
AATATTGTGT CCCACAACCA AGGAGCCATG CCACTCAAAG GTTCCACAAC CTGGAACAC 300  
AAAGATTCCA GAGCCAGGCT GTACCAAGGT CCCTGAGCCA GGCTGTACCA AGGTCCCTGA 360  
GCCAGGCTGT ACCAAGGTCC CTGAGCCAGG TTGTACCAAG GTCCCTGAGC CAGGCTGTAC 420  
CAAGGTCCCT GAGCCAGGTT GTACCAAGGT CCCTGAGCCA GGCTACACCA AGGTCCCTGA 480  
ACCAGGCAGC ATCAAGGTCC CTGACCAAGG CTTTCATCAAG TTTCTGAGC CAGGTGCCAT 540  
CAAGTTCCT GAGCAAGGAT ACACCAAGT TCCTGTGCCA GGCTACACAA AGCTACCAGA 600  
GCCATGTCTT TCAACGGTCA CTCCAGGCCC AGCTCAGCAG AAGACCAAGC AGAAGTAATT 660  
TGGTGCACAG ACAAGCCCTT GAGAAGCCAA CCACCAGATG CTGGACACCC TCTTCCCATC 720  
TGTTTCTGTG TCTTAATGT CTGTAGACCT TGTAAATCAGC ACATTGTGAC CCCAAGCCAT 780  
AGTCTCTCTC TTATTGTGAT CCTAAAATA CGTACTATAA AGCTTTGTT CACACACACT 840  
CTGAAGAATC CTGTAAGCCC CTGAATTAA GAGAAAGTCT TCATGGCTTT TCTGGTCTTC 900  
GGCTGCTCAG GGTTCATCTG AAGATTGGA TGAAGAGAAA TGCAATGTTT CTGCTCTTCC 960  
CTCATTAAT TGCTTTTAAT TECA

Seq ID NO: 46 Protein sequence:  
Protein Accession #: NP\_005407.1

1 11 21 31 41 51  
MSSYQKQKTF TPPPQLQQQQ VKQPSQPPFP EIFVPTTKEP CHSKVPQPGN TKIPEPGCTK 60  
VPEPGCTKVP EPQCTKVPEP GCTKVPEPGC TKVPEPGCTK VPEPGYTKVP EPGSIKVPDQ 120  
GFIKFPPEGA IKVPEQGYTK VPVPGYTKLP EPCPSTVTPG PAQKQTKQK

Seq ID NO: 47 DNA sequence  
Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51  
5 GCGTCGTGTG CAGGCGTCCC CGGGCTGTGG ATAATTAGAC ACGTTCCTCC CTCATTGCC 60  
AAGGCTCGTT AGAATTCGCC CTAGAGCTGT ATCATGTATT TTCTTCAAA TTAACCTTGC 120  
TTGCAATTAA GCTTAGGGAA CCAGCAACAA AAGCAAACTT GGCCCGAGGT CGTTCACGCG 180  
GAAATATGAT TAGAGAACT TCTTCCCGA TTTAAGGGGA AAGATTCTCT CGGCCAGCGC 240  
TTTGGGGAAG TCGCCCGGAC CGCAGAGGCG ACGACAGGGG AGCAGGAAGC TGCTCAGCGT 300  
AGTCCGCGTT GCGCGCAGCG GTGGCCTTCC TCATCTGGGC GATGTGGGCT CCTAGAAGAG 360  
10 TAAGATAAC ATCTTGAAGA TGACTTCTGT ACGGTTTGAG CCCAACTGCA CACTCATGAC 420  
TTGGAGCTGC CCGTGGAGAT TACAGTTTAC CAAACACATT CATGAACATA ATCTCATTTA 480  
CTAAAAACTT TGTGAGAATT TTCTTTTACT AAAATTTTTT CTTATTACAA A

Seq ID NO: 48 DNA sequence:

Nucleic Acid Accession #: CAT cluster

15 1 11 21 31 41 51  
20 TTCCAAATTT TTTTITTTGT AATAAGAAAA AATTTTAGTA AAGAAAAATT CTCACAAAGT 60  
TTTGTAGTAA TGAGATTATG TTTCTGAATG TGTITGGTAA ACTGTAACTC CACAGGGCAG 120  
CTCCAAGTCA TGAGTGTGCA GTTGGGCTCA AACCGTACAG AAGTCATTTC CAGGATGTTA 180  
TCTTACTCTC TCTCGGAGCC CACATCGCCC AGATGAGGAA GGCCACCGCT GCCGCCAAGC 240  
CGGACTACCG TGAGCAGCTT CCGTCTCCCC TGTCGTGCCC TCTGGGTTGG GGGCACTTTC 300  
25 CCCAAAGCGC TGGCCGCAAG AATCTTTCCC CTTAAATCGG GGAAGAAGTT TCTCTAATCC 360  
ATTTTGGCGG TGAACGACCT CGGGCCAAAG TTGCTTTTGT TGCTGTTTCC CTAAGCTTAA 420  
TTGCAAGCRA AGTTAATTTG AAAGAAAAATA CATGATACAG CTCTAGGGCG AATTCTAAGC 480  
AGCCTTGGGC AATGAGGGAA GAACGTGTCT AGTTATCCAC AGCCCGGGGA CGCTGCACA 540  
CGACGCT

Seq ID NO: 49 DNA sequence

Nucleic Acid Accession #: CAT cluster

30 1 11 21 31 41 51  
35 TCTTCTCTCT GCTGCTCGTT TGCTCTCTCT GTGCTCTTCT TCTTCTTTC CTOGCGGCT 60  
CTGCGGACC TCGTGTGTCT CTCTCTGAT GCGGGGGGGG GGGAGAAGCT GACCGGTGAG 120  
ACCGTAGACC CGAAACCAAT GGGTGTGACA AGCCGGTCCG CGGCTTTTTT GGGAGAAGCC 180  
GACACATGCA GACCACTTTT CCGTGAACNG CATGACCATG TTATTACTAT GGGCCGCTCT 240  
40 CCCAACCAAA GTGTTTAAAA CTTTITAGGG CACCCCAAAA ATTTTTTTTT TTTTITTTTT 300  
TTCAATTAAA AAACCTTAAT ATTTATATTA AATACAAAGA TACCCAAACC CTTTATGCTT 360  
CTTCTCTGTA TCGTGTCTCT TTTTCTTGA CAGCATCTCC AITTTTTTTC TGCTGCTTCA 420  
TCGCTGTAGC CATGGGAATC CGTTTCATTA TTATGTTAGC AATATGAGT GCTGTATTCC 480  
TAAAGAAACT GACACAGGAG AATCACTTGA ACTTGGGAGG CAGAGTTTGC AGTGAGCCGA 540  
GATTGAACCA GTGCATCTCA GCCTTGGCAG CGGAGCAAGA TTCTGTGACA GTTCTGAAG 600  
45 TGCTGTATC GTCTGTCAGC CCCATCTCTG GTTCCATTGC GCTGCCAGGC AGGGTGTCTG 660  
GAGTGGGGA GAGCTGGTCT ATATATCCGG GTGAAGCTCA GCTGTGGCAC ACCTTGGATG 720  
CGGAGTCTCT CCGGCCCCCG GGGACCTAGT ATTTTGGCCA CGAGGTGACA CCAACAAAG 780  
GAGACAGCAT CATTTATGAG CCTGCAGCAT CCACCTTACT GCTGTATCCA GTTTCATTG 840  
ACTG

Seq ID NO: 50 DNA sequence

Nucleic Acid Accession #: L05187

Coding sequence: 1991..2260

55 1 11 21 31 41 51  
60 CTGCAGGGAG GCAGGTAGAA AAGGCTTTTG GGTITTCAGG TGGGGGGCAG TCTAGCCTGA 60  
TCAGAAAGGA GGAAGAGGCC AGGGCAGATG TCTGGGTGGA GTGAAGGGAA AAAGTGATCC 120  
CAGAAGAAGG ATTAGCCCTT GAAAGTCCCT GAAGTAGGAG AAGGGTAAAG GTGTGGTTGG 180  
TGAAGGAAGG CAGGTTTCC CAGATTAGCA ACCAGTCAGG GGGAGGAAGG TGAGAGTGGG 240  
AGATCATATA GTAAATTAT CTGAATGTGT GTAGTTTAAAT GGAATTGGGA AAAAGATGGG 300  
GGAATGGAT GGAAGGTCTT GGAATCTGAG ACAAGGGGTC TATAATCAGT CCATTTTCATT 360  
ATTTCTAGCT TCCACCTTCA CCAAGGCAGA CAAGGAGGGC CCACCTCAGC TCCTCTGCTC 420  
65 CCCCTCCCTT TCCACCTAT TCATGTGTGC AAGAGTSCCC TGTCCCACAG AACACGGGGA 480  
ACAACCATCT CAATGACAAG GACAGCAAGT GGCAGGCTC AACAGGACTC AGATGTCCCC 540  
CCAGGGTTAA CTCATGAAAC CCTCCATGAA GCGCTGTGCT CACCCCTCCC TCAAGGCAAG 600  
CCCTGCACCT GGGTCTGAGG ATGAGGGTGG CAGTGAATAA TAGGCCAGTG ACATCATTTT 660  
CAGCCAGCTA GTGCTCAAAA ATATCAGGTG GTGTTCATCA AATAAGCCGA GCCAACCCGT 720  
GATGAGGATG GTAGTGTGAG TCATGTGTGA CAGGTGAGGA ATGAAAACAG AGTGCCCGAG 780  
70 AGCTTCTATT TCCTTAGGCG AGGGCTCAT CATCTTATAA AAGCCAGCTG GCCATTGCCT 840  
TCACACCAAA CCCAAGGGAC CACACAGCCC ATTCTGCTCC GTATACCAGG TAAGTCTCTG 900  
ATTGCAACAA ACTGGCAATT CTAGTGTACT TTTTCATTAT TAGAAATTAG CTAAGGCAAG 960  
ATATGTGTAA GCAGGTTTTC CCAGGTTTTC AATGGGAGAT AGAGATAGT GGAATATCTT 1020  
TATTTTAAGT TAAATTACAG TCTGGATTG AAAGGACCTT AGAGATGGTT AGGGCTCCCA 1080  
75 CCTCAGTAGA TAGTCATTGA ACTGGGAGTC CTGAGAAGA TTGTTCAAAT GCCCATGGGA 1140  
AGTTCATAGC AGAAGTAGAA CTCAGGCCAG AGCACTCTCA GTAACACTGC AATTTCCCCC 1200  
TGACAAGATA TTTATAGAAA TTTTAATTTA TTAGATGGAT CTCTACTGAG CATTTATTCC 1260  
ATTTAAGGCA GTATGCTAGG CACTTTGGAC AAATCAATGC CCTAACGTAC TTACTTAACA 1320  
AACAATAAAC CTAGCAGGAA GGTAAATCAT ATATATAAAT AAATGAAATG CAAAGTAGAT 1380  
80 AGTAATTGGC ATGACGGAGA TGGGCAGAGA AGGGCTGTGC ACTTTTGGGA GACTTGTCTA 1440  
AGGAGACCTC TAGGGTGTCA AGTGATGTGA GCTATGTGG AGGGGTATTG GGAACAGCAG 1500  
AGATGGGAAG AAAAGCATTT GGAAGGGACT GTGTAAGCAC AGACCAGAAG CAAACACATA 1560  
GAGGCTTAGA TGAATATAAA GCCATCCTAT AAGTCACAGG CTTTCTACAT GGTACTAGGA 1620  
GAGGAAGAGT GTCTGATGCC ATTTTCCAAA AGACCTAATA TGCGGACCTC ATGTCCCTCA 1680  
85 GAAGCCAGCT TTAGTAGGCG ATTTTCCAG AACAGATATA AGGTGCTCTG GGTAGGAAGG 1740  
GAGCCAGAA GAGAACTCCA ATAAATGGA GCAGAAGAAA TTGCCTTTTA GTCCTCCTC 1800  
TTCAAGGGC CTGAAATTA TCCAAGCTTA TTTTATTTTT AAATGTAATG GGGAGCTAA 1860

GGGAGATGAA AGGCTTTCTC TTCTAAAGGG TCCTGAAATA AAATCTGTTT GGCATTGAAT 1920  
TTGTATCCAT CTTTCTTTAA TTGAATCACT GTGTGAGCTT TCTGTCTCTA GAAAAAACA 1980  
CATTGGAAGC ATGAATTTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA 2040  
GCAGCAGCAG GTGAACAACAC CTTGCCAGCC TCCACCCAG GAACCATGCA TCCCAAAAC 2100  
CAAGGAGCCC TGCCAAACCA AGGTGCTGTA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG 2160  
CCAGCCCAGG ATTCCAGAGC CCTGCCAGCC CAAGGTGCCT GAGCCCTGCC CTTCAACGGT 2220  
CACTCCAGCA CCAGCCCAGC AGAAGACCAA GCAGAAGTAA TGTGGTCCAC AGCCATGCCC 2280  
TTGAGGAGCT GGCCTCTGGA TACTGAACAC CCTACTCCAT TCTGCTTATG AATCCCATTT 2340  
GCCTATTGAC CCTGCAGTTA GCATGCTGTC ACCCTGAATC ATAATCGCTC CTTTGCACCT 2400  
CTAAAAAGAT GTCCCTTACC CTCATTCTGG AGGCTCCTGA GCCTCTGCGT AAGGCTGAAC 2460  
GTCTCACTGA CTGAGCTAGT CTCTCTGTTG CTCGGGTGCA TTTGAGGATG GATTGCGGGA 2520  
AGGTCAGTG ACCATCCCTA G

Seq ID NO: 51 Protein sequence:  
Protein Accession #: AAC26838

1 11 21 31 41 51  
MNSQQQKQPC TPPPOPOQQQ VKQPCPPPPQ EPCIPKTKEP QPKVPEPCPH PKVPEPCQPK 60  
IPEPCQPKVP EPCPSTVIPA PAQKRTKQK

Seq ID NO: 52 DNA sequence  
Nucleic Acid Accession #: NM\_002638.1  
Coding sequence: 120-473

1 11 21 31 41 51  
CAATACAGCT AAGGAATTAT CCCTTGTAAG TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60  
GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120  
TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGACG CTGGTTCTAG 180  
AGGAGCTGT CACGGGAGTT CTGTGTAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240  
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300  
CGCAAGAGCC AGTCAAAGGT CCACTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA 360  
TCCGTTGGGC CATGTTGAAT CCCCTAACC GCTGCTTGA AGATACTGAC TGCCCAAGAA 420  
TCAAGAAATG CTGTGAAGGC TCTGCGGGA TGGCCTGTTT CGTCCCCAG TGAAGGGAGC 480  
CGGTCTTGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540  
TGTGCTCTT CCCCTTCCA CACTGTCCAT TCTTCTCCC ATTGAGGATG CCCAGCGCTG 600  
GAGCTGCCCT TCTCATCCAC TTTCCAATAA A

Seq ID NO: 53 Protein sequence:  
Protein Accession #: NP\_002629.1

1 11 21 31 41 51  
MRASSFLIVV VFLLIAGTLVL EAAVTGVFVK GQDTVKGKRP FNGQDFVKQ VSVKQDKVK 60  
AQEPVKGPVS TRPGSCPIL IRCAMLNPPN RCLKDTCPG IKKCCGSGC MACFVPQ

Seq ID NO: 54 DNA sequence  
Nucleic Acid Accession #: NM\_019618  
Coding sequence: 75-584

1 11 21 31 41 51  
GGCAGAGGCC AGGATTCAGT CCCCTGGACT GTAGATAAAG ACCCTTCTT GCCAGGTGCT 60  
GAGACAACCA CACTATGAGA GGCCTCCAG GAGACGCTGA TGGTGGAGGA AGGGCCGTCT 120  
ATCAATCAAT GTGTAAACCT ATTACTGGGA CTATTAATGA TTGGAATCAG CAAGTGTGGA 180  
CCCTTCAGGG TCAGAACCTT GTGGCAGTTC CACGAAAGTGA CAGTGTGACC CCAGTCACTG 240  
TGTCTGTTAT CACATGCAAG TATCCAGAGG CTCTTGAGCA AGGCAGAGGG GATCCCATTT 300  
ATTGGGAAT CCAGATCCA GAAATGTGTT TGTATTGTGA GAAGGTGGA GAACAGCCCA 360  
CATTGCAGCT AAAAGAGCAG AAGATCATGG ATCTGTATGG CCAACCCGAG CCCGTGAAAC 420  
CCTTCTCTTT CTACCGTGCC AAGACTGGTA GGACCTCCAC CCTTGAGTCT GTGGCCTTCC 480  
CGGACTGGTT CATTGCTCTC TCCAAGAGAG ACCAGCCCAT CATTCTGACT TCAGAACTTG 540  
GGAAGTCATA CAACACTGCC TTTGAATTAA ATATAATGA CTGAACTCAG CCTAGAGGTG 600  
GCAGCTGGT CTTGTCTTA AAGTTCTGG TTCCCAATGT GTTTTGTCT ACATTTTCTT 660  
AGTGTCAATT TCACGCTGGT GCTGAGACAG GGGCAAGGCT GCTGTTATCA TCTCATTTTA 720  
TAATGAAGAA GAAGCAATTA CTTCATAGCA ACTGAAGAAC AGGATGTGGC CTCAGAAGCA 780  
GGAGAGCTGG GTGGTATAAG GCTGCTCTCT CAAGCTGGTG CTGTGTAGGC CACAAGGCAT 840  
CTGCATGAGT GACTTTAAGA CTCAAAGACC AAACACTGAG CTTTCTTCTA GGGGTGGGTA 900  
TGAAGATGCT TCAGAGCTCA TGCGGTTAC CCACGATGGC ATGACTAGCA CAGAGCTGAT 960  
CTCTGTTTCT GTTTTGTCTT ATTCCCTCTT GGGATGATAT CATCCAGTCT TTATATGTTG 1020  
CCAATATACC TCATTGTGTG TAATAGAACC TTCTTAGCAT TAAGACCTTG TAAACAAAAA 1080  
TAATTCTTGT GTTAAGTTAA ATCATTTTTG TCCTAATTGT AATGTGTAAT CTTAAAGTTA 1140  
AATAAACTTT GTGTATTAT ATAATAAAAA AAAAAAAAAA AAA

Seq ID NO: 55 Protein sequence:  
Protein Accession #: NP\_062564

1 11 21 31 41 51  
MRGTPGDADG GGRAVYQSMC KPITGTINDL NQQVNTLQGG NLVAVPRSDS VTPVTVAVIT 60  
CKPYEALEQG RGDPIYLGIO NPENCLYCEK VGEQPTLQLK EQKINDLYGQ PEPVKPFLFY 120  
RAKTRGTSTL ESVAPPDWFI ASSKRDQPII LTSELGKSYN TAFELNIND

Seq ID NO: 56 DNA sequence  
Nucleic Acid Accession #: NM\_003125  
Coding sequence: 65-334

1 11 21 31 41 51  
AGCAGTTCTA AGGGACCATA CAGAGTATTC CTCTCTTCAC ACCAGGACCA GCCACTGTTG 60  
CAGCATGAGT TCCAGCAGC AGAAGCAGCC CTGCATCCCA CCCCTCAGC TTCAGCAGCA 120  
GCAGGTGAAA CAGCCTTGCC AGCCTCCACC TCAGGAACCA TGCACTCCCA AAACCAAGGA 180  
GCCCTGCCAC CCCAAGGTGC CTGAGCCCTG CCACCCCAAA GTGCCCTGAGC CCTGCCAGCC 240  
CAAGCTTCCA GAGCCATGCC ACCCCAAGGT GCCTGAGCCC TGCCCTTCAA TAGTCACTCC 300  
AGCACGAGCC CAGCAGAAGA CCAAGCAGAA GTAATGTGGT CCACAGCCAT GCCCTTGAGG 360  
AGCGGCCAC CAGATGCTGA ATCCCTATC CCATCTGTG TATGAGTCCC ATTTGCTCTG 420  
CAATTAGCAT TCTGTCTCCC CCAAAAAGA ATGTGCTATG AAGCTTTCTT TCCTACACAC 480  
TCTGAGTCTC TGAATGAAGC TGAAGGTCTT AGTACCAGAG CTAGTTTTCA GCTGCTCAGA 540  
ATTCACTCTG AGAGAGACTT AAGATGAAAG CAAATGATTG AGCTCCCTTA TACCCCATTT 600  
AAATTCACCT TCAATTCCA

Seq ID NO: 57 Protein sequence:  
Protein Accession #: NP\_003116

1 11 21 31 41 51  
MSSQQQKQPC IPPPQLQQQ VKQPCQPPPO EPCIPKTKEP CHPKVPPEPCH PKVPEPCQPK 60  
LPEPCHPKVP EPCPSIVTPA PAQKTKQK

Seq ID NO: 58 DNA sequence  
Nucleic Acid Accession #: NM\_001793.2  
Coding sequence: 71-2560

1 11 21 31 41 51  
AAAGGGGCAA GAGCTGAGCG GAACACCCGGC CCGCCGTGCG GGCAGCTGCT TCACCCCTCT 60  
CTCTGAGGCC ATGGGGCTCC CTCGTGGACC TCTCGGTCT CTCTCCTTC TCCAGGTTTG 120  
CTGGCTGCGAG TCGCGGGCCT CCGAGCCGTG CCGGGCGGTG TTCAGGGAGG CTGAAGTGAC 180  
CTTGAGAGCG GAGAGCGCGG AGCAGGAGCC CGGCCAGCGG CTGGGGAAGG TATTATGAGG 240  
CTGCCCTGGG CAGAGCCAG CTCTGTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300  
TGCGGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360  
ATCCAAAGCT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420  
TGAAATATGC AAGGTTCCTT TCCCCAGAG ACTGAATCAG CTCAGTCTA ATAAAGATAG 480  
AGACACCAAG ATTTTCTACA GCATCAGCGG GCGCGGGSCA GACAGCCCCC CTGAGGGTGT 540  
CTTCGCTGTA GAGAAGGAGA CAGGCTGGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600  
GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTGAGAG AATGGTGCTT CAGTGGAGGA 660  
CCCCATGAAC ATCTCCATCA TCGTGACCGA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720  
GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCTTACCA GGTACTTCTG TGTGTCAGGT 780  
GACAGCCACG GATGAGGATG ATGCCATCTA CACCTACAAT GGGGTGGTGG CTTACTCCAT 840  
CCATAGCCAA GAACCAAGG ACCCAACGA CCTCATGTTT ACCATTCAAC GGAGCACAGG 900  
CACCATCAGC GTCACTCCA GTGGCCTGGA CCGGGAAGAA GTCCCTGAGT ACACACTGAC 960  
CATCCAGGCC ACAGACATGG ATGGGAGCGG CTCACCAACC ACGGCAGTGG CAGTAGTGGG 1020  
GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080  
GCCTGAGAAAT GCATGGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140  
CAACTCACC GGTGGCGTGG CCACCTACCT TATCATGGGC GGTGACGAGG GGGACCATTT 1200  
TACCATCACC ACCCACTCTG AGAGCAACCA GGGCATCTCT ACAACGAGGA AGGGTTTGGG 1260  
TTTTGAGGCC AAAAACCGA ACACCTGTA CGTTGAAGTG ACCAACGAGG CCCCTTTTGT 1320  
GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380  
ACCTGTGTTT GTCCCACTCT CCAAGTGTGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440  
GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAGGAG AATCAAAAGA TCAGCTACCG 1500  
CATCCTGAGA GACCCAGCAG GGTGGCTAGC CATGAGCCCA GACAGTGGGC AGGTCAACG 1560  
TGTGGGCACC CTGACCGCTG AGGATGAGCA GTTTGTGAGG AACACATCT ATGAAGTCA 1620  
GGTCTGGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGTAAAC 1680  
ACTGATTGAT GTCAATGACC ATGGCCCACT CCCTGAGCCC CGTCAGATCA CCATCTGCAA 1740  
CCAAAGCCCT GTGCGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800  
CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGAAGCGCAG AGGTCAACGA 1860  
GGAAGGTGAC ACAGTGTCTT TGTCCCTGAA GAAGTTCTCT AAGCAGGATA CATATGACGT 1920  
GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980  
GTGCGACTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGAAGGGAG GTTTCATCCT 2040  
CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGCTGCTGC TTTTGTGTGT 2100  
GAGAAAGAG CCGAAGATCA AGGAGCCCTT CCTACTCCCA GAAGATGACA CCCGTGACAA 2160  
CGTCTTCTAC TATGGCGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220  
GCTCCACCGA GGTCTGGAGG CAGGCGCGGA GGTGGTTCTC CGCAATGAGG TGGCACCAC 2280  
CATCATCCCG ACACCCATGT ACGTCTCTCG GCCAGCCAAC CCAGATGAAA TCGGCAACTT 2340  
TATAATTGAG AACCTGAAGG CCGCTAACAC AGACCCCAAC GCGCCGCGCT TCCCTGAGCT CCTCACCTC 2400  
CTTGGTGTTC GACTATGAGG GCAGCGGCTC CGACGCGCGG TCCCTGAGCT CCTCACCTC 2460  
CTCCGCTCC GACCAAGACC AAGATTACGA TTATCTGAAC GAGTGGGGCA GCGCTTCAA 2520  
GAAGCTGGCA GACATGTACG GTGGCGGGGA GGAAGCTAGG GCGGCTGTCC TGCAGGGCTG 2580  
GGGACCAAAC GTGAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640  
GACTTGGGAG CTGTGTCAGG AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700  
ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTGTACTTC 2760  
AGCACTGAAA ACCCTTCCAC CTGGGCCAGG GTTGCCTCAG AGGCCAAGTT TCCAGAAGCC 2820  
TCTTACTGCG CGTAAATGCG TCAACCTGTG GTCTCTGGCC TGGGCTGTCT GTGACTGACC 2880  
TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGA ACTTAATTTT 2940  
TTTTTTTAAAT GCTATCTTCA AAACGTAGA GAAAGTTCTT CAAAGTGCA GCCCAGAGCT 3000  
GCTGGGCCCC CTGGCGTCTC TGCAATTTCT GTTTCAGAC CCAATGCTC CCATTCGGA 3060  
TGGATCTCTG CGTTTATATA CTGAGTGTGC CTAGGTGGCC CCTTATTTTT TATTTTCCCT 3120  
GTTGCGTTCG TATAGATGAA GGGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTTTAT 3180  
TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 59 Protein sequence:



Protein Accession #: NP\_001784.2

1 11 21 31 41 51  
5 MGLPRGGLAS LLLLQVCWLQ CAASEFCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG 60  
QEPALFSDN DDPFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120  
KGFPPQRLNQ LKSNKDRDTK IFYSITGPGA DSEPPGVFAV EKETGWLLEN KPLDRESIAK 180  
YELFGEHASE NGASVEDPMN ISIIIVTDQND HKPKFTQDTF RGSVLEGLVP GTSMVMQVTAT 240  
10 DEDDAIITYN GVVAYSISHSQ EPKDPHDLMP TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300  
TMDGDGSGTT TAVAVVEILD ANDNAPMFPD QKYEAHVPEV AVGEHVQRLT VTDLDAPNSP 360  
AWRATYILMG GDDGDHFTIT THPEBNQGIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420  
PTSTATIVVH VEDVNEAPVF VPSPKVEVVO EGIPTGEFVC VYTAEDPDKE NQKISYRILR 480  
DPAGNLAMP DSGQVTAUGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLTLID 540  
15 VNDHGFVPEP RQITICNQSP VRQVLNITDK DLSPTSPFQ AQLTDDSDIY WTAEVNEEGD 600  
TVVLSLKLFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFILPVL 660  
GAVLALLFL LVLALLVRKK RKIKPELFLP EDDTRDNVFP YGEEGGGEED QDYDITQLHR 720  
GLEARPEVVL RNDVAPTIIIP TMYRPRPAN PDEIGNFIIE NLKAANTDPT APFYDTLLVF 780  
DYEAGSGDAA SLSSLTSSAS DQDQDYDYLN EWGSRPKKLA DMYGGGEDD

Seq ID NO: 60 DNA sequence  
Nucleic Acid Accession #: Eos sequence.  
Coding sequence: 162-428

1 11 21 31 41 51  
25 GCGTTCGGTT GCGCGCGGAT TCGAACGTTT GGACTGAGGT TTTCTGCCT GAAGAAGCGT 60  
CATACGGACC GGATGTGTTT CGCTGGCCCA GTGTCCCGG AGCTGTGTG CGATACAGAG 120  
AGCACTCGG AGCTGAGGCC AGCTGGTACT TGACAGAGAG GATGGCGCTG TCGACCATAG 180  
30 TCTCCAGAG GAAGCAGATA AAGCGGAAGG CTCCCGTGG CTTTCTAAG CGAGTCTTCA 240  
AGCGAAGAA GCCTCAACTT CGTCTGGAGA AAGTGGTGA CTATTGGTC CATCTGAAC 300  
GTTTACTGTT TGTTCATCGA TTAGCAGAAG AGTCCAGGAC AAGCGCTGT GCGAGTAAAT 360  
GTAGAGTCAT TAACAGGAG CATGTACTGG CGCAGCAAA GGTAATTCTA AAGAAGAGCA 420  
GAGGTTAGAA GTCAAGAAC ATATTCTTGA AAGTTATGAT GCATTCCTTT GGGTGGTAAC 480  
35 AGATCATAAA GACATTTTTT ACACATCAGT TAATATGGGA TTATTAAATA TTGG

Seq ID NO: 61 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
40 MALSTIVSQR KQIKRKAPRG FLKRVFKRKK PQLRLKSGD LLVHLNCLLF VHLAEESRT 60  
NACASKCRVI NKEHVLAAAK VILKKSRI

Seq ID NO: 62 DNA sequence  
Nucleic Acid Accession #: NM\_000094.2  
Coding sequence: 99-8933

1 11 21 31 41 51  
50 GGGCTGGAGG GGGCTGGGCT TCGGACCTGC CAAGGCCACC GCAGGGGGA GCAAGGGACA 60  
GAGGGGGGGG TCCTAGCTGA CGGCTTTTAC TGCTAGGAT GACGCTGCGG CTTCTGGTGG 120  
CCGCGCTCTG CCGCGGGATC CTGGCAGAGG CGCCCGGAGT GCGAGCCAG CACAGGGAGA 180  
GAGTGACCTG CAGCGGCTTT TACGCGCTG ACATTGTGTT CTTACTGGAT GGCTCCTCAT 240  
55 CATTGGGCGG CAGCAATTTT CGCGAGGTCC GCAGCTTTCT CGAAGGGCTG GTGCTGCCTT 300  
TCTCTGGAGC AGCCAGTGCA CAGGGGTGTC GCTTTGCCAC AGTGCACTAC AGCGATGACC 360  
CAGGACAGA GTTCGGCCTG GATGCACTTG GCTCTGGGGG TGATGTGATC CGGCCCATCC 420  
GTGAGCTTAG CTACAAGGGG GGCAACACTC GCACAGGGGC TGCAATTCTC CATGTGGCTG 480  
ACCATGTCTT CTTGCCCCAG CTGGGCGGAC CTGGTGTCCC CAAGGTCTGC ATCCTGATCA 540  
60 CAGACGGGAA GTCCAGGAG CTGGTGGACA CAGCTGCCCA AAGGCTGAAG GGGCAGGGGG 600  
TCAAGCTATT TGCTGTGGGG ATCAAGAAATG CTGACCTTGA GGAGCTGAAG AGAGTTGCTT 660  
CACAGCCAACT CTTCCGACTT TTCTTCTTGG TCAATGACTT CAGCATCTTG AGGACACTAC 720  
TGCCCTCTGT TTCCCGGAGA GTGTGACAGA CTGCTGGTGG CGTGCTGTG ACCCGACCTC 780  
CGGATGACTC GACCTCTGCT CCAAGAGACC TGGTGTGCTC TGAGCCAAGC AGCCAATCCT 840  
65 TGAGAGTACA GTGGACAGCG GCCAGTGGCC CTGTGACTGG CTACAAGGTC CAGTACACTC 900  
CTCTGACGGG GCTGGGACAG CCACTGCGGA GTGAGCGGCA GGAGGTGAAC GTCCCACTG 960  
GTGAGACCAG TGTGCGGCTG CGGGTCTTCC GGCCACTGAC CGAGTACCAA GTGACTGTGA 1020  
TTGCCCTCTA CGCCAAACAG ATCGGGGAGG CTGTGAGCGG GACAGCTCGG ACCACTGCCC 1080  
TAGAAGGGCC GGAACCTGAC ATCCAGATA CCAAGGCCA CAGCCTCTGT GTGGCTTGGC 1140  
70 GGAGTGTGCC AGGTGCCACT GGCTACCGTG TGACATGGCG GGTCTCTCAGT GGTGGGCCCA 1200  
CACAGCAGCA GGAGCTGGCG CTTGGGAGG GTTCAGTGT GCTGCGTGAC TTGGAGCCTG 1260  
GCACGGACTA TGAGGTGACC GTGAGCACCC TATTGGCCG CAGTGTGGGG CCGGCCACTT 1320  
CCCTGATGGC TGCACTGAC GCTTCTGTTG AGCAGACCTT CGCCCGGCTC ATCCTGGGCC 1380  
CCACATCCAT CCTCTTTTCC TGGAACTTGG TGCTGAGGC CGTGGCTAC CGGTTGGAAT 1440  
75 GCGCGCGTGA GACTGGCTTG GAGCCACCGC AGAAGGTGCT ACTGCCCTCT GATGTGACCC 1500  
GCTACCAATT GGATGGGCTG CAGCCGGGCA CTGAGTACCG CCTCACACTC TACACTCTGC 1560  
TGGAGGGCCA CGAGGTGGCC ACCCTGCAA CGGTGGTTCC CACTGACCA GAGCTGCTG 1620  
TGAGCCCTGT AACAGACCTG CAAGCCACCG AGCTGCCCGG GCAGCGGGT CGAGTGTCTT 1680  
GGAGCCAGT CCCTGTGGCC ACCAGTACC GCATCATTGT GCGCAGCACC CAGGGGGTTG 1740  
80 AGCGGACCTT GGTGCTTCTT GGGAGTCAGA CAGCAITCGA CTTGGATGAC GTTCAGGCTG 1800  
GGCTTAGCTA CACTGTGCGG GTGTCTGCTC GAGTGGGTTC CCGTGAAGGC AGTGCCAGTG 1860  
TCTCAGTGT CCGCGGGAG CCGGAAACTC CACTGTGCTT TCCAGGGCTG CGGGTGTG 1920  
TGTGAGTGC AACCGAGTG AGGGTGGCCT GGGGACCGT CCCTGAGCC AGTGGATTTC 1980  
GGATTAGCTG GAGCACAGGC AGTGTGCGG AGTCCAGCCA GACACTGCCC CCAGACTCTA 2040  
CTGCCACAGA CATCACAGGG CTGACGCTG GAACCACTA CCAGGTGGCT GTGTGGTAC 2100  
85 TGCGAGGCG CAGCAGAGGG CCTGCTGCAG TCATCGTGGC TCGAACCGCA CCACTGGGCC 2160  
CAGTGAGGAC GGTCCATGTG ACTCAGGCCA CGAGCTCATC TGTCAACATT ACCTGGACCA 2220  
GGGTCTCTGG CGCCACAGGA TACAGGGTTT CTGTGCACTC AGCCACGGC CCAGAGAAAT 2280

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

CCCAGTTGGT TTCTGGGGAG GCCACGGTGG CTGAGCTGGA TGGACTGGAG CCAGATACTG 2340  
 AGTATACGGT GCATGTGAGG GCCCATGTGG CTGGCTGGA TGGGCCCCCT GCCTCTGTGG 2400  
 TTGTGAGGAC TGCCCCGTGAG CTTGTGGGTC GTGTGTGAG GCTGCAGATC CTCAATGCTT 2460  
 CCAGGCGAGT TCTAGGATC ACCTGGGTAG GGGTCACTGG AGCCACAGCT TACAGACTGG 2520  
 CCTGGGGCCG GAGTGAAGGC GGCCCCATGA GGCACACAGT ACTCCAGGA AACACAGACT 2580  
 CTGCAAGAT CCGGGTCTC GAAAGTGGAG TCAGCTACTC AGTGCAGTGT ACTGCACCTG 2640  
 TCGGGGACCG CGAGGGCACA CTTGTCTCCA TTGTGTGCAC TACGCCGCTT GAGGCTCCGC 2700  
 CAGCCCTGGG GAGCCTTCCG GTGGTGACGC GGGGGGAGCA CTGCTGAGG CTGGCTGGG 2760  
 AGCCGGTGGC CAGAGCGCAG GGCCTCCTTC TGCACTGGCA ACCTGAGGGT GGCCAGGAAC 2820  
 AGTCCCGGGT CTTGGGGCCC GAGCTCAGCA GCTATCACTT GGACGGGCTG GAGCCAGCGA 2880  
 CACAGTACCG GTGAGGCTG AGTGTCTTAG GGCCTGGTGG AGAAGGGCCC TCTGCAGAGG 2940  
 TGAATGCGCG CACTGAGTCA CCTGTGTTC CAAGCATTGA ACTAGTGTG GTGACACCT 3000  
 CGATCGACTC GTGTGCTTGG GCTTGGACTC CAGTGTCCAG GGCATCCAGC TACATCTTAT 3060  
 CTTGGCGGCC ACTCAGAGGC CTTGGCCAGG AAGTGCCTGG GTCCCGCAG ACACITCCAG 3120  
 GGATCTCAAG CTCCACGGG GTGACAGGGC TAGAGCTTGG CGTCTCTTAC ATCTTCTCCC 3180  
 TGACGCTGTG CTTGAGTGTG GTGGGGTTC CTGAGGCATC TGTACACAG ACGCCAGTGT 3240  
 GCCCCCTTGG CTTGGCGGAT GTGGTGTTC TACCACATGC CACTCAAGAC AATGCTCACC 3300  
 GTGGGAGGGC TAGAGGAGG GTCTGGAGC GTCTGGTGT GGCACITGGG CTTCTTGGG 3360  
 CACAGGCACT TCAGGTGGC CTGCTGTCTT ACAGTCATCG GCCCTCCCCA CTGTTCACAC 3420  
 TGAATGGCTC CCAATGACCTT GGCATTATCT TGCAAGAGAT CGGTGACATG CCTACATGG 3480  
 ACCCAAGTGG GAACAACTGT GGCACAGCG TGGTCAACG TCACAGATAC ATGTGGCAC 3540  
 CAGATGCTCC TGGGCGCCGC CAGCAGTAC CAGGGGTGAT GGTCTGCTA GTGGATGAAC 3600  
 CCTTGAGAGG TGACATATTC AGCCCCATCC GTGAGGCCCA GGCCTTCTGG CTTAATGTGG 3660  
 TGATGTGGG AATGGCTGGA GCGGACCCAG AGCAGCTGG TCGCTTGGG CCGGTATGG 3720  
 ACTCTGTCCA GACCTTCTTC GCGGTGATG ATGGGCCAAG CTTGGACAGC GCAGTCAGT 3780  
 GTCTGGCAC AGCCCTGTGT CAGGCATCCT TCACTACTCA GCCCGGGCCA GAGCCCTGCC 3840  
 CAGTGTATTG TCCAAAGGGC CAGAAGGGG AACCTGGAGA GATGGGCTG AGAGGACAA 3900  
 TTGGGCTTCC TGGCGACCTT GGCCTCCCGG GCAGGACCGG TGCTCCCGG CCGCAGGGG 3960  
 CCGCTGGAAG TGCCACTGCC AAGGGCGAGA GGGGCTTCCC TGGAGCAGT GGCCTCCAG 4020  
 GCAGCCCTGG CCGCGCCGGG AATCCTGGGA CCCCTGGAGC CCTTGGCCTA AAGGGCTCT 4080  
 CAGGGTGGC AGCCCTGTGT GGGGACCGG GAGAGCGAGG ACCTCGAGG CCAAGGGGG 4140  
 AGCCGGGGG TCCCGGACAA GTCATCGAG GTGAAGACC TGGGCTTCTT GGGCGGAAAG 4200  
 GGGACCTGG ACCATCGGGC CCCCCTGGAC CTGCTGGACC ACTGGGGGAC CCAGGACCCC 4260  
 GTGGCCCCC AGGGCTTCTT GGAACAGCCA TGAAGGTGA CAAAGCGAT GGTGGGAGC 4320  
 GGGGTCCCC TGGACAGGT GAAGGTGGCA TTGCTCTGG GGAGCTGGG CTGCGGGTCT 4380  
 TTCCGGGAG TGCCGACCCC CAAGGCCCGG TTGGCCCCC TGGAAAGAAA GAGAGAAAAG 4440  
 GTGACTTGA GATGTGAGCT CAGGCCCTCC CAGGACAACC TGGGTCTCCG GGTGAGCAGG 4500  
 GCGCAAGGG ACCCTCTGGA GCTATTGGCC CCAAAGGTGA CCGGGCTTT CAGGGCCCC 4560  
 TGGGTGAGGC TGGAGAGAG GCGGAACGTG GACCCCCAGG CCGACCGGGA TCCCGGGGG 4620  
 TGCCAGGGGT TGTGTGAGCT CCGTGGACCA AGGCTCTGA AGGGCCACCA GGACCCACTG 4680  
 GCGCCAAAG AGAGAAGGG GAGCCTGGTC GCGCTGGGGA CCGCTGAGTG GTGGGACTG 4740  
 CTGTGTGCTG ACCCAAGGGA GAAAGGGGAG ATGTGGGGCC CGCTGGGGCC AGAGGAGCTA 4800  
 CCGGAGTCCA AGGGGAACGG GCGCCACCCG GCTTGGTCT TCTGGAGAC CTTGGCCCCA 4860  
 AGGGAGACCC TGGAGACCGG GGTCCCATG GCTTACTGG CAGAGCAGGA CCGCCAGGTG 4920  
 ACTCAGGGCC TCTTGGAGAG AAGGGAGACC CTGGGCGGCC TGGCCCCCA GACCTGTGT 4980  
 GCGCCGAGG AGCAGATGTT GAAGTTGGAG AGAAAGGTGA CAGGGTCTCT CCGGTGACC 5040  
 CCGGTGTGCC TGGAAAGCA GCGCAGGCTG GCTTCCGGG GGCACCTGGA GTTCCGGGGC 5100  
 CTGTGGGTGA AAAGGGAGC CAGGAGATC CTGGAGAGGA TGGACGAAAT GGCAGCCCTG 5160  
 GATCATCTGG ACCCAAGGTT GACCGTGGGG AGCCGGGTCC CCGAGGACCC CCGGGACGGC 5220  
 TGGTAGACAC AGGACCTGGA GCCAGAGAGA AGGGAGAGCC TGGGACCGG GACAAAGAG 5280  
 GTCTCGAGG CCGCAAGGTT GATCTGGCC TCCCTGGAGC CCGTGGGAA AGGGCATTG 5340  
 AAGGGTTTCG GGGACCCCCA GCGCCACAG GGGACCCAGG TGTCCGAGGC CCAGCAGGAG 5400  
 AAAAGGGTGA CCGGGTCTCC CTTGGGCTGG ATGGCCGGAG CCGACTGGAT GGGAAACAG 5460  
 GAGCCGCTGG GCGCTCTGG CCGAATGGTG CTGAGGCAA AGCTGGGGAC CCGGGAGAG 5520  
 ACCGGCTTCC AGGCCCTCGT GGAAGAACAG GCTTCTCTGG CCGCTCTGGT CCGCTGGAT 5580  
 TACCGGAAA GCGCAGGCGAG GATGGGAAAC CTGGCTGAA TGGAAAAAC GGAGAACCTG 5640  
 GGGACCTGG AGAAGACGG AGGAAGGGAG AGAAAGGAGA TTCAGGCGCC TCTGGGAGAG 5700  
 AAGGTGTGA TGGCCCCAAG GGTGAGCGTG GAGCTCTGG TATCTTGA TATCTTGA 5760  
 CTCCAGGCT CCGAGGGCCA GTGGGCCCTC CTGGCCAGGG TTTTCTGTGT GTCCAGGAG 5820  
 GCACGGGGCC CAAGGTTGAC CTTGGGAGA CTGGATCCA AGGGAGCAG GGCCTCCCTG 5880  
 GAGAGCGTGG CTTGCGAGGA GAGCCTGGAA GTGTGCGGA TGTGTATCGG TTTCTGAAA 5940  
 CTGTGCGCAT CAAGGCATCT GCGCTGCGG AGATCGTGA GACCTGGAT GAGAGCTCTG 6000  
 GTAGCTTCTT GCTGTGCCC GAACGGCGTC GAGGCCCAA GGGGACTCA GCGAACAGG 6060  
 GCGCCCGAG CAAGGAGGGC CCAATCGCTT TTTCTGGAGA ACCCGGGCTG AAGGGCGACC 6120  
 GTGGAGACCC TGGCCTCTAG GGGCCACTG GTCTGGCCCT TGGGAGAGG GCGCCCGCG 6180  
 GGCCTTCGG CTTTGGCGG GAGCCTGGAA AGCCTGTAT TCCCGGGCTC CAGGCGAGG 6240  
 CTGGGGTGT GGGAGAGGCA GGAAGGCCAG GAGAGAGGG AGAAGCGGGA GAGAAAGGAG 6300  
 AACGTGAGA ACAGGGCAGA GATGGCCCTC CTGGACTTCC TGGAAACCCCT GGGCCCCCG 6360  
 GACCCCTGG CCGCAAGGTG TCTGTGATG AGCCAGGTCC TGGACTCTCT GGAAGACAGG 6420  
 GACCCCTGG ACTCAAGGTT GCTAAGGGGG AGCCGGGCG CAATGTGTAC CAAGGTCCA 6480  
 AAGGAGACAG GGGTGTCCA GGCATCAAAG GAGACCGGG AGAGCCTGGA CCGAGGGGTC 6540  
 AGGACGGCAA CCGGGTCTA CCGAGAGAGC GTGGTATGG TGGGCTGAA GGAAGCGCG 6600  
 GTCTGCAGG TCCAGAGAGC CCGCTGGCC CAGTGGGTGG TCTGAGAGC CTTGAGCAC 6660  
 CTGTGCCCC GGTCTTGTG GGCCTGCGAG GACCCCAAG ACCTTCTGGC CTGAGGGGG 6720  
 AGCCTGGAGA GACAGGACT CAGGACCGG GCTTACTGG ACCTACTGGA GCTGTGGGAC 6780  
 TTCTTGACC CCGCGCCCT TCGAGCCCTT TGGTCCCA GGGGTCTCA GGTTTGCTG 6840  
 GACAAGTGGG GAGACAGGG AAGCCGGGAG CCGCAGTCG AGATGGTGGC AGTGAAGAG 6900  
 ATGGAGACAG AGGGAGCCCT GGTGTGCCAG GGTACACAG TCTGCTGGC CTTGTGGAG 6960  
 CTAAGGAGA ACCTGGCCCC ACGGGGGCC CTGGACAGG TGTGTGCGG CTCCCTGGAG 7020  
 CAAAGGAGA GAAAGGAGCC CTTGAGGCC TTGCTGGAGA CTTGTGGGT GAGCCGGAG 7080  
 CCAAGGTGA CCGAGGACT CAGGGCGCG GAGGCGAGAA GGGTGAAGCT GGCCTGCGAG 7140  
 GGGAGCCCG AGACCTTGG GAAGATGTC AGAAAGGGG TCCAGGACCC AAAGGTTTCA 7200  
 AGGGTACCC AGGAGTCCG GTCCCGGCT CCGCTGGGCC TCTTGGCCCT CAGGTGTGA 7260  
 AGGAGATCT GGGCCTCCTT GGCCTGGCG GTGCTCTGG TGTGTGGG TTCCCGGGTC 7320  
 AGACAGGCC TCGAGGAGAG ATGGGTCAAG CAGGCCCTAG TGGAGAGCG GGTCTGGCAG 7380  
 GCGCCCGAG GAGGAAGGA ATCCAGGAC CCGTGGGCC ACCTGAGCA CCGGGTCTAG 7440  
 TGGGACACC TGGGCGCTCT GGCATCAAAG GAGACAAGG AGACCTGGA GTAGGGCTGC 7500

	CTGGGCCCGG	AGGCGAGCGT	GGGAGGCCAG	GCATCCGGGG	TGAAGATGGC	CGCCCCGGCC	7560
	AGGAGGGACC	CCGAGGACTC	ACGGGGCCCC	CTGGCAGCAG	GGGAGAGCGT	GGGAGAGAGG	7620
	GTGATGTTGG	GAGTGCAGGA	CTAAGGGGTG	ACAAGGAGAG	CTCAGCTGTG	ATCCTGGGGC	7680
5	CTCCAGGCCC	ACGGGGTGCC	AGGGGGGACA	TGGGTGAACG	AGGGCCTCGG	GGCTTGGATG	7740
	GTGACAAAGG	ACCTCGGGGA	GACAAATGGG	ACCTGTGTGA	CAAGGGCAGC	AAGGGAGAGC	7800
	CTGGTGACAA	GGGCTCAGCC	GGGTGCCAG	GACTGCGTGG	ACTCCTGGGA	CCCCAGGGTC	7860
	AACCTGGTGC	AGCAGGAGTC	CCTGTGACC	CGGGATCCCC	AGGAAAGGAT	GGAGTGCCCTG	7920
	GTATCCGAGG	AGAAAAAGGA	GATGTTGGCT	TCATGGGTCC	CGGGGGCCTC	AAGGGTGAAC	7980
10	GGGGAGTGAA	GGGAGCCTGT	GGCCTTGATG	GAGAGAAGGG	AGACAAGGGA	GAAGCTGGTC	8040
	CCCCAGGCCC	CCCCGGGCTG	GCAGGACACA	AAGGAGAGAT	GGGGGAGCCT	GGTGTGCCGG	8100
	GCCAGTCGGG	AGGGCCTGCG	AAGGAGGGCC	TGATCGGTCC	CAAGGGTGAC	CGAGGCTTTG	8160
	ACGGGCAGCG	AGGCCCCAAG	GGTGACCAGG	GCGAGAAAGG	GGAGCGGGGA	ACCCAGGAAA	8220
	TGGGGGGCTT	CCGAGGCCCC	AGTGAATAAG	ATGGCTCTGC	TGGTCCCCCA	GGGCCACCTG	8280
15	GCAGTGTGG	TCCAGAGGCG	CCCGAAGGAC	TTCAGGGCCA	GAAGGGTGAG	CGAGGTCCCC	8340
	CCGGAGAGAG	AGTGGTGGGG	GCTCCTGGGG	TCCTCGGAGC	TCCTGGCGAG	AGAGGGGAGC	8400
	AGGGGCGGCG	AGGGCCTGCC	GGTCTCGAG	GCGAGAAGGG	AGAAGCTGCA	CTGACGGAGG	8460
	ATGACATCCG	GGGCTTTGTG	CGCCAAAGAG	TGAGTCAGCA	CTGTGCTCTG	CAGGGCCAGT	8520
	TCATGCAATC	TGGATCAGCA	CCCTCCCTTA	GTTATGCTGC	AGACACTGCC	GGCTCCAGC	8580
20	TCCATGCTGT	GCCTGTGCTC	CGCTCTCTCT	ATGCAGAGGA	GGAAGAGCGG	GTAACCCCTG	8640
	AGGATGATGA	GTACTCTGAA	TACTCCGAGT	ATTCTGTGGA	GGAGTACCAG	GACCCCTGAAG	8700
	CTCCTTGGGA	TAGTATGATC	CCCTGTTCCC	TGCCACTGGA	TGAGGGCTCC	TGCATCGCTT	8760
	ACACCCCTCG	CTGGTACCAT	CGGGCTGTGA	CAGGCAGCAC	AGAGGCTCTG	CACCCCTTTG	8820
	TCATGTTGG	CTGTGGAGGG	AATGCCAAC	GTTTGGGAC	CGGTGAGGCC	TGCGAGCGCC	8880
25	GCTGCCACCC	CCGGTGTGTC	CAGAGCCAGG	GGACAGGTAC	TGCCAGGAC	TGAGGCCAG	8940
	ATAATGAGCT	GAGATTACAG	ATCCCTCGGA	GGAGTCGGGG	TCTCAGCAGA	ACCCCACTGT	9000
	CCCTCCCTCT	GGTGGCTAGAG	GCTTGTGTGC	ACGTGAGCGT	GCGAGTGAC	GTCCGTATT	9060
	TCACTGACTT	GGTCCCGTGG	GTCTAGCCTT	CCCCCTGTG	GACAAACCCC	CATTGTGGCT	9120
	CTGCGCACCC	TGGCAGATGA	CTCACTGTGG	GGGGTGGCT	GTGGGCAATG	AGCGGATGTG	9180
30	ACTGGCGTCT	GACCCGCCCC	TTGACCCAAG	CCTGTGATGA	CATGTGCTGT	ATTCTGGGGG	9240
	GCATTAAGC	TGCTGTTTTA	AAAGGCAAAA	AA			

Seq ID NO: 63 Protein sequence:

Protein Accession #: NP\_000085.1

35	1	11	21	31	41	51	
	MTLRLVLAAL	CAGILAEAPR	VRAQHERERV	CTRLYAADIV	FLLDGSSSIG	RSNFREVRSP	60
	LESLVLPFSG	AASAGGVIFA	TVQYSDDPRT	EPGLDALGSG	GDVIRAIREL	SYKGNTRTG	120
40	AAILLVADHV	FLPOLARPGV	PKVCILITDG	KSQLDVTAA	QRLKGQGVKL	FAVGKINADP	180
	EELKRVASQP	TSDFFFFVND	FSILRLTLLPL	VSRRVCTTAG	GVPVTRPFDD	STSAPRDLVL	240
	SEPSQSRLRV	QMTAASGPVT	GKVKQYTPLT	GLGQPLPSER	QSVNVPAGET	SVRLRGLRPL	300
	TEYQVTVIAL	YANSIGEAUS	GTARTTALEG	PELTITQNTA	HSLLVAMRSV	PGATGYRVTV	360
	RVLSSGGTQQ	QELGPGQSGT	LLRDLPEPGTD	YEVTVSTLFG	RSVGPATSLM	ARTDASVEQT	420
45	LRPVILGPTS	ILLSNWLYPE	ARGYRLEWRR	ETGLEPPQKV	VLPSTDVTRYQ	LDLGLQPGTEY	480
	RLTLYTLLEG	HEVATPATVW	PTGPPLPVSP	VTDLQATELP	GQRVRVWSP	VPGATQYRII	540
	VRSTQGVERT	LVLPGSQTAF	DLDDVQAGLS	YTVRVSAVVG	PREGSASVLT	VRRPETPLA	600
	VPGLRVVSD	ATRVRAVWGP	VPGASGFRIS	WSTGSGPSSS	QTLPPDSTAT	DITGLQPGTT	660
	YQVAVSVLRG	REEGPAIVIV	ARTDPLGPRV	TVHVTQASSS	SVTITWTRVP	GATGYRVSWH	720
50	SARGHEKSQL	VSGEATVAEL	DGLEPDTFTT	VHVRHVAVAG	DGPPASVVVR	TAPEPVGRVS	780
	RLQILNASSD	VLRIITWVGVT	GATAYRLANG	RSEGGPMRHQ	ILPGNTDSAE	IRGLEGGVSY	840
	SVRVTAIVGD	RSQTPVSIIV	TTTPEAPPAL	GTLHVVRQGE	HSLRLRWEPV	PRAQGFLLHW	900
	QPGGQGEQSR	VLPGLSSSYH	LDGLEPATQY	VRLLSVLGA	GEGPSAEVTA	RTESFVPVSI	960
	ELRVVDTSID	SVTLAWTPVS	RASSYILSWR	PLRGPQGEVP	GSPQTLPGIS	SSQRTVGLPE	1020
55	GVSYIFSLTP	VLDGVRGPEA	SVTQTFVCPR	GLADVVLPH	ATQDNAHRAE	ATRVLERLV	1080
	LALGPLGPOA	VQVGLLSYSH	RPSPLFPLNG	SHDLGILQ	IRDMFYMDPS	GNNLSTAVVT	1140
	AHRYMLAPDA	PGRRHQHVPV	MVLLVDEPLR	GDIFSPFIRE	QASGLNVVML	MGAGADPEQL	1200
	RRLAPGMDSV	QTFFAVDGPP	SLDQAVSGLA	TALCQASFTT	QPRPEPCPVY	CPKQKGEPE	1260
	EMGLRGQVGP	PGDPLGPGRT	GAPGPGGPPG	SATARGERG	PGADGRPGSP	GRAGNPGTGP	1320
60	APGLRGSPGL	PGPRGDPGER	GPRGPKGEPG	APGQVIGGEG	PGLPGRKGD	GPSPGPGPRG	1380
	PLGDPGPRGP	PGLPGTAMKG	DKGDRGERGP	PGPGGGIAP	GEPGLPGLPG	SPGPGGVVPG	1440
	PGKKGKGDSD	EDGAPGLPGQ	PGSPGEQGPR	GPPGAIGPKG	DRGFPGLGE	AGEKGERGPP	1500
	GPAGSRGLPG	VAGRPGAKGP	EGPPGPTGRQ	GEKGEPRPG	DPAVVGPAVA	GPKEKGDVG	1560
	PAGPRGATGV	QGERGPPGLV	LPDGPMPKGD	PGDRGPGLT	GRAGPPGDSG	PPGEKGDPR	1620
65	PGPPGPGVGR	GRDGEVGEKG	DEGPPGDPGL	PGKAGERGLR	GAPGVRGPGV	EKGDDQDPGE	1680
	DGRNGSPGSS	GPKGDREGE	PPGPPGRLVD	TGPGAREKGE	PGDRGQEGPR	GPKGDGGLPG	1740
	APGERGIBGF	RGPFGPGQDP	GVRGPAGEKG	DRGPPGLDGR	SLDGLKPGAA	GPSPGFGAAG	1800
	KAGDPGRDGL	PGLRGEQGLP	GPSGPPGLPG	KPGEDGKPLG	NGKNGECPDP	GEDGRKGEKG	1860
70	DSGASGREGR	DGPKGERGAP	GILGPGGPPG	LPGPVGPPGQ	GFPVGPVGGT	PKGDRGETGS	1920
	KGEQGLPGER	GLRGEPSGVP	NVDRLLTAG	IKASALREIV	ETWDESSGSP	LPVPERRRGP	1980
	KGDSGEQGPP	GKEGPIGPGV	ERGLKGDGRD	PGPGGPPGLA	LGERGPPGPS	GLAGEPGKPG	2040
	IPGLPGRAGG	VGEAGRPGER	GERGEKGERG	EQGRDGPPLG	POTPGPPGPP	GPVSVDEPG	2100
	PGLSGEQGPP	GLKGAKEGEP	SNGDQGPKGD	RGVPGIKGDR	GEPGPRQDQ	NPGLPGERGM	2160
	AGPEGKPLGQ	GPRGPPGPGV	GHDGPGPPGA	PGLAGPAGPQ	GPSGLKGEPE	ETGPPGRGLT	2220
75	GPTGAVGLPG	PPGPGSLVGP	QGSPLPGQV	GETGKPGAPG	RDGASGKDG	RQSPGVGSP	2280
	GLPGVPVPGK	EPGPTGAPG	AVVGLPGAAG	EKGAPGGLAG	DLVGEPGAAG	DRGLPGPRGE	2340
	KGEAGRAGEP	GDPGEDGQPK	APGPKGPKGD	PGVGVGSPFG	PPGPPGVKGD	LGLPGLPGAP	2400
	GVVGPFGQTG	PRGEMGQPGP	SGERGLAGPP	GREGIPGLAG	PPGPPGSPVG	PGASGLKGD	2460
	GDPGVGLPGP	RGERGEPGIR	GEDGRPGQEG	PRGLTGPPGS	RGERGEKGDV	GSAGLKCDKG	2520
80	DSAVILGPPG	PRQAGKDMGE	RQPRGLDGD	GPRGNDGDPG	DKGSKGEPGD	KGSAGLPGLR	2580
	GLLGPQGGPG	AAGIPDPGDS	PGKDGVPGIR	GEKGDVGFGR	PRGLKGERGV	KGACGLDGEK	2640
	GDKGEAGPPG	RPGLAGHKEG	MGEPPGVPGS	GAPGKEGLIG	PKGDRGFDGQ	PGPKGDQGEK	2700
	GERGTPGIGG	FPGPSGNDGS	AGPPGPGSV	GPRGPEGLQG	OKGERGPPE	RVVGAPGVPG	2760
	APGERGEQGR	PGPAGPRGEK	GEAALTEDDI	RGFVRQEMSQ	HCACQGGPIA	SGSRFLPSYA	2820
85	ADTAGSQLHA	VPVLRVSHAE	EEERVPPEDD	EYSEYSEYSV	EYQDPEAPIN	DSDDPCSLEP	2880
	DEGSCTATTL	RNYHRAVTGS	TEACHPFVYG	GCGGNANRFG	TREACERRCP	PRVVQSQGTG	2940
	TAQD						

5  
1 11 21 31 41 51  
ATGTCCTTATC AACACGACGA GTGCAAGCAG CCTGCCAGC CACCTCCTGT GTGCCCCACG 60  
CCAAAGTGCC CAGAGCCATG TCCACCCCG AAGTGCCCTG AGCCCTGCC ACCACCAAAG 120  
10 TGTCACACAG CCTGCCACC TCAGCAGTGC CAGCAGAAT ATCTCCTGT GACACCTTCC 180  
CCACCCTGCC AGCCAAAGTA TCCACGAAGA AGCAGATAA

15            1                    11                    21                    31                    41                    51  
              |                    |                    |                    |                    |  
MSYQQQCKQ PCQPPFVCPT PKCEPCPPK KCCEPCPPK CPQPCPQQC QQKYPVFTS 60  
PCQPKYPPK SK

	1	11	21	31	41	51	
25	TAGTCGGAGC	GAGGTGGCGA	GTGCTGAGC	CCGCGCGCGC	CCCGAGAGCG	GCTGCAGCCG	60
	CCGCGCGCGG	GAAGGAGAGG	GGGAGGCGCG	CCCGAGCGCG	CGCGCGCGCG	GCACCGCGCG	120
	CCGCGCGCCAC	CCAGCCGACC	GGAGTCCGCG	GCCAGCCGCG	CAGCCTCCGC	GGGCCCCGCG	180
30	CGGGGCGGGG	GGCGCGGGCC	ACAGGCCCCT	CGGCGCGCGG	TGGTTTGAG	ACCCTGGGCG	240
	CGGATGTGCG	CCGCGCGCGG	TTAGGATGAG	TCTCGGGTGC	GGGAGGAGAG	CGCGCGAGCC	300
	CGCGCGCGCC	GAGCGCGCGG	CAGAGAGCTC	GCGCGCGCGC	CGCGCGCGCG	CGCGCGCGCG	360
	GGCGGGCCCC	GACGCGCGCC	GCGCGCGCCC	GGGCGCCCGA	CACACATGAG	ATTCTTCAGG	420
	CTCATCTTCA	AGTGCTTCGT	GGACTGCTTC	TGACTTGGCC	GCGCGCGCCC	CGCACCCGCG	480
35	CGTCCGCGCG	CGCGCGCGCT	CCCGCGCGCG	GCGCGCCCCC	GGCCCCCGCG	CGCGCGCGCG	540
	CCTCGGGGCG	CTCCCCGGTG	CCCGCGGTGC	CCCCCGCGCT	ACCGCGCGCG	CCCGTGAGGC	600
	GGCGCGACCC	CGCGCGGGCC	GTGGCGCGCG	CGCGGGCGCT	GGGGAAGAGC	ACCGCGAGAG	660
	ADGGCATCTA	TAGCTGTGTC	GGCGACGAGA	TAGAGGGCCC	CCTCATCGCG	CCCGGGCGCG	720
	AGGGGGCCCC	GGCCAAAGGG	GACGGCCCCG	GGGCGCTGGG	GACAACCGCG	GGCGCGCTGG	780
40	CGGTGCGGCC	GGCGGAGACC	TGGACGCGCC	AGATGGACTT	CATCATGTCT	TGGCTGGGCT	840
	TGCGCGTGGG	CTTGGGCAAC	GTGTGGGGCT	TCCCTTACCT	GTGCTACAG	AACCGCGGAG	900
	GTGTGTTCTT	TATTCCCTAT	CTCTGATGCT	CCCTGTGTGG	AGGAATCCCG	ATTTCTTCTT	960
	TAGAGATCTC	CTGGGGGACG	TTTCATGAAG	CGGCGAGCAT	CAATGTCTGG	AACATCTCTC	1020
	CCCTGTTCAA	AGGCTCTGGC	TAGGCTTCCA	TCTGTATGCT	CTTCTACTGC	AACACCTACT	1080
45	ACATCATGGT	GCTGGCTGCG	GGCTTCTATT	ACCTGTGTAA	GTCTTTTACC	ACACGCTGTC	1140
	CCTGGGCGAC	ATGTGGCCAC	ACCTGGGAAC	CTCCGAGACT	CGTGGAGATC	TTCCGCGATG	1200
	AGAGCTGTGC	ATATGCGAAC	CTGGCGCAAC	TCACCTGTGA	CGGATGTGCT	GACCGCCGGT	1260
	CCCTGTGATC	CGAGTCTGCG	GAGAACCAAG	TCTTGAGGCT	GTCTGGGGAG	CTGGAGGTCT	1320
	CGGGGGCCCT	CAACTGGGTC	GTGACCTTCT	GTCTGTGGCG	CTGTGGGGTG	CTGGTCTACT	1380
50	TCTGTGCTCG	GAAGGGGGTC	AAATCCAGCG	GAAGATGTGT	GTACTTCACT	GTCATCACTC	1440
	CTCATGCTGG	CTGTGTGCTG	CTGTGTGGTG	GTGGAGTGGC	GCTCGGCGGC	GGCCGGATGT	1500
	GCATCATTTA	CTATCTCAAG	CCCTAGTGTG	CAAAAGCTGG	GTCCCTCAG	GTGTGGATAG	1560
	ATGCGGGGAG	CCGATTTTCT	TTTTTCTAOC	CACTTGGCCCT	GGGGGCCCTC	ACAGCCCTGG	1620
	CGAGCTACAA	CGAGTTTCAAC	AACAACTGTG	CAAGAAGGCG	CATCATCTGC	GCTCTCATCA	1680
55	ACAGTGGGAC	CAGCTTCTTT	GCTGGCTTGC	TGGTCTTCTC	CATCTGGGCG	TTCATGGGTC	1740
	CAGAGCAGGG	GGTCGACATC	TCCAAGGTGG	CAGAGTGGAG	CGCGGGGCTG	GCTCTCATCG	1800
	CCTACCGCGC	CCTGTCTCAG	CTGATGCCAG	TGGCCCCACT	CTGGGCTGCC	CTGTCTTCTT	1860
	TGATGCTTGG	GCTGCTTGGG	CTGACGAGCC	AGTTTGTAGG	TGTGGAGGGC	TTTATCAGCG	1920
	GCTCTCTGTA	CCCTCTCCGT	GCCTCTACTG	ACTTTCGTTT	CCAAAGAGGAG	ATCTCTGTGG	1980
60	CCCTCTGTGT	TGCCCTCTGC	TTTGTCTATG	ATCTCTCCAT	GGTGACGTAT	GGCGGGATGT	2040
	AGCTGTTTCA	GCTGTTTGAC	TACTACTGCG	CACGGGCGCG	CACCTGCTCT	TGGCAGGCTC	2100
	TTTGGGATGT	CGTGGTGGTG	GCCTGGGTGT	ACGGAGCTGA	CCGCTTCAAT	GACGACATGT	2160
	CTGTGATGAT	CGGGTACCGA	CTCTGGCCCT	GGATGAAATG	GTGCTGGTCC	TTCTTCAACC	2220
	CGCTGGTGTG	CATGGGCACT	TTTCATTTCA	CGGTTGTGTA	CTCGGCTTCA	CTGGTCTACA	2280
65	ACAAACACTA	CGTGTACCCG	TGGTGGGGTG	AGGGCACTGG	CTGGGCTCTC	GGCTCTGTCT	2340
	CCATGCTGTG	CGTCCGCTCG	CACCTCTGCG	GCTGGCTTCT	CAGGGCCAGG	GAGCAGATGG	2400
	CTGAGCGCTG	GCAGCACCTG	ACCCAGCCCC	TCTGGGGCTC	CCACCACTCT	GAGTACCGAG	2460
	CTCAGGACCG	AGATGTCAAG	GGGCTGACCA	CCCTGACCCC	AGTGTCCGAG	AGCAGCAGGG	2520
	TGCTGTGGTG	GGAGAGTGTG	ATGTGACAA	CACGCTTACA	TACCAGGCTC	ACCTCTGGTA	2580
70	GCATAGGACG	CCCTGTCTTC	AGGCCCGACG	CAACCCCTCA	GGGGGCTCTG	CTTCTCCGTA	2640
	CCCTTTTGGG	GCTGCTGTGG	GGGAGGAGGG	GAGAAAGGAC	CATGAGTGTCT	CATTAAACAA	2700
	ACTTTTTTCA	TTTTTAATAA	AACGCCAAAA	ATATACACAC	CCACCAAAAA	TAGATGCTCT	2760
	TCCCCCTCTCA	GGCCTAGCCG	AGCTGGTCTT	AGGCCCGGCC	TAGTGCACCA	CCCCCACCAC	2820
	CAGTGTCTCA	CGTCTCTCTG	CCCTGGCACG	CCACCCCTCT	GCCCACTCTT	CCAGGCTCTG	2880
75	CTGCGCAGCA	CACCCGTGGG	TGACCCCTCA	CCCGAGAGAG	AGCAGTGAGC	GCTTGGGAAA	2940
	TGTGAGGAGG	GGAAGGAGGG	AGAGACGGGA	GGGAGGAGAG	AGAGGAGAAG	GAGGACGAGG	3000
	GAGGGGCGAG						

CTTCTGTGTA GCAGCTTTAA CCCACGTTTG TCTGTCACTG CCAGTCCCGA GACGCGTAG 3780  
 TGACCCCAAG AAAGGCTTCC CCGACACCCA GACAGAGGCT GCAGGGCTGG GGCTGGGTGA 3840  
 GGTGGCGGG CCTGGCGGGA CATTCTACTG TGCTAAAAAG CCACTGCAGA CATAGCAATA 3900  
 AAAACATGTC ATTTTCC

Seq ID NO: 67 Protein sequence:  
 Protein Accession #: NP\_005620.1

1 11 21 31 41 51  
 MAKSAENGI YSVSGDEKKG FLIAPGPDGA PAKGDGPVGL GTPGGRLAVP PRETWTRQMD 60  
 PIMSCVGPFAV GLGNVVRFFY LCYKNGGGVF LIPYVLIALV GGIPIPFLEI SLGQFMKAGS 120  
 INWNICPLF KGLGYASWVI VFYCNYYIM VLAWGFYILV KSFTTTLFWA TCGHTWNTPD 180  
 CVEIFRHEDC ANASLANLTC DQLADRRSPV IEFWENKVLK LSGGLEVPFA LNWEVTLCLL 240  
 ACWVLVYFCV WKGKVSSTGKI VYFTATFPYV VLVVLLVRGV LLPGALDGI IYLPKPDWSKL 300  
 GSPQVNIIDAG TQIFPSYAIG LGALTALGSY NRPNNNCYKD AIIILALINSQ TSFFAGVVPV 360  
 SILGFMAEBQ GVHISKVRES GPGLAFIAYP RAVTLMFPVAP LNAALFFPML LLLGLDSQFV 420  
 GVEGFITGLL DLLPASYFYR FQRRIISVALC CALCFVIDLS MVTDCGMVYP QLFDYYSASG 480  
 TTLNQAQFWE CVVVAMVYGA DRFMDIACM IGYRCPWMK WCVSFFTPLV CMGIFIFNVV 540  
 YYEPLVYNNT YVYPWNGEAM GWAPALSSML CVPLHLLGCL LRKAGTMABR WQHLTQPIWG 600  
 LHFLEYRAQD ADVRLGLTLT PVSESSKVVV VESVM

Seq ID NO: 68 DNA sequence  
 Nucleic Acid Accession #: NM\_021953.1  
 Coding sequence: 178-2469

1 11 21 31 41 51  
 GGCACGAGGG GGACCCGCGC GGTCCGGCGC GAGCCCCCGT CCGGGGCCCT GGCTCGGCC 60  
 CCAGGTTGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGCG GCGGCGACTG 120  
 CAGTCTGGAG GGTCCACACT TGTGATCTCT AATGGAGAGT GAAAACGCGA ATTCATAATG 180  
 AAAGCTAGCC CCGCTCGGCC ACTGATCTCT AAAAGACGGA GGCTGCCCTT TCCTGTTCAA 240  
 AATGCCCCAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGAGTCT 300  
 AATCAAGCAG AGGCCTCCAA GGAAGTGGCG GAGTCCAAC CTTCGCAAGT TCCAGCTGGG 360  
 ATCAAGATTA TTAACACCC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCACAAT 420  
 GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAGG GAAAAGAGAG TGGCAGTAGT 480  
 GGCCCAACA AATTCTCTCT CATCAGCTGT GGGGGAGGCC CAACTCAGCC TCCAGGACTC 540  
 CGGCTCAAA CCAAAACAG CTATGATGCC AAAAGACAG AAGTGACCTT GGAGACCTTG 600  
 GGACCAAAAC CTGCAGCTAG GGATGTGAAT CTCTCTAGAC CACCTGGAGC CCTTTGCGAG 660  
 CAGAAACCGG AGACCTGTGC AGATGGTGAG GCAGCAGGCT GCACTATCAA CAATAGCCTA 720  
 TCCAACATCC AGTGCTCTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780  
 CAAGAGATGG AGGAAAGGGA GAATTGTCACT CTGGAGCAGC GACAGGTAA GGTGTAGGAG 840  
 CCTTCGAGAC CATCAGCGTC CTGGCAGAAC TCTGTGTCTG AGCGGCCACC CTACTCTTAC 900  
 ATGGCCATGA TACAATTCGC CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960  
 ATCTATACGT GAGTTGAGGA CCACCTTCCC TACTTTAAGC ACATTCGCAA GCCAGGCTGG 1020  
 AAGAACTCCA TCCGCCACAA CCTTTCCTCG CACGACATGT TTGTCCGGGA GACCTCTGCC 1080  
 AATGGCAGG TCTCTCTCTG GACCATTCAC CCGAGTGCCA ACCGCTACTT GACATTTGAC 1140  
 CAGGTGTTTA AGCCACTGGA CCCAGGGTCT CCACAATTCG CCGAGCACTT GGAATCACAG 1200  
 CAGAAACGAC CGAATCCAGA GCTCCGCGCG AACATGACCA TCAAAACGGA ACTCCCCCTG 1260  
 GCGGCACCGG GGAAGATGAA GCCACTGCTA CCACGGGTCA GCTCATACCT GGTACCTATC 1320  
 CAGTTCCCGG TGAACCAAGT ACTGTTGTTG CAGCCCTCGG TGAAGGTGCC ATTTGCCCTG 1380  
 GCGGCTTCCC TCATGAGCTC AGAGCTTGCC CGCCATAGCA AGCGAGTCCG CATTTGCCCT 1440  
 AAGGTGTGTC TAGCTGAGGA GGGGATAGCT CCTCTTCTCT CTGCAGGACC AGGGAAAGAG 1500  
 GAGAAACTCC TGTTTGAGGA AGGGTTTCTT CCTTTGCTTC CAGTTGAGAC TATCAAGGAG 1560  
 GAAGAAATCC AGCTGGGGGA GGAATGCGA CACTTAGCGA GACCCATCAA AGTGGAGAGT 1620  
 CCTCCCTTGG AAGAGTGGCC CTCCCGCGCC CCATCTTTCA AAGAGGAATC ATCTCACTCC 1680  
 TGGGAGGATT CGTCCCAATC TCCACCCCA AGACCCAAGA AGTCTACAG TGGGCTTAGG 1740  
 TCCCAACCCC GGTGTGCTC GGAATGCTT GTGATTCAAC ACAGGGAGAG GAGGGAGAGG 1800  
 AGCCGCTCTC GGAGGAACAC GCATCTACTG CCTCCCTGTG TGGATGAGCC GGAGCTGCTC 1860  
 TTCTCAGAGG GGCACAGTAC TTCCCGCTGG GCGCAGAGC TCCGTTTCCC AGCAGACTCC 1920  
 TCTGACCTCG CTCTCCAGCT CAGCTACTCC CAGGAAGTGG GAGGACCTTT TAAGACACCC 1980  
 ATTAAGGAAA CGCTGCCCAT CTCTCCACCC CCGAGCAAAT CTGTCTCCCC CAGAACCCCT 2040  
 GAATCCTGGA GGCTCAGGCC CCCAGCCAAA GTAGGGGGAC TGGATTTCAG CCCAGTACAA 2100  
 ACCTCCAGG GTGCTCTGTA CCCCTTGCTT GACCCCTGGG GGCTGATGGA TCTCAGCACC 2160  
 ACTCCCTTGC AAAGTGCTCC CCCCCTTGAA TCACCGCAAA GGCTCCTCAG TTCAGAACCC 2220  
 TTAGACCTCA TCTCGTCCC CTTTGGCAAC TCTTCTCCCT CAGATATAGA CGTCCCCAAG 2280  
 CCAGGCTCCC CCGAGCCACA GGTTCCTGGC CTTGCAGCCA ATCGTTCTCT GACAGAAGGC 2340  
 CTGGTCTCGG ACACAATGAA TGACAGCTCT AGCAAGATCC TGCTGGACAT CAGCTTTCTT 2400  
 GGCTTGGAAG AGGACCACT GGGCCCTGAC AACATCAACT GGTCCAGTT TATTCTCTAG 2460  
 CTACAGTAGA GCCCTGCCCT TGCCCTGTGT CTCAGCTGT CCACCATCCC GGGCACTCCA 2520  
 AGGCTCAGTG CACCCAGAGC CTCTGAGTGA GGCAGCAGG CAGGGACTGT TCTGCTCTCT 2580  
 ATAGCTCCCT GATGCTCTGAT TATGCAAAAG TAGCAGTCAC ACCCTAGCCA CTGCTGGGAC 2640  
 CTTGTGTTC CCAAGAGTAT CTGATTOCTC TGCTGTCCCT GCCAGGAGCT GAAGGGTGGG 2700  
 AACAAACAA GCAATGGTGA AAAGAGATTA GGAACCCCCC AGCCTGTTTC CATTCTCTGC 2760  
 CCAGCAGTCT CTACCTTCCC CTGATCTTTG CAGGTGTGTC CGTGTAAATA GTATAAATTC 2820  
 TCCAAATTAT CCTTAATTA TAAATGTAAG CTTATTTCCT TAGATCATTA TCCAGAGACT 2880  
 GCCAGAAGT GGGTAGGATG ACCTGGGGTT TCAATTGACT TCTGTTCTCT GCTTTTAGTT 2940  
 TTGATAGAAG GGAAGACCTG CAGTGACCGG TTTCTTCCAG GCTGAGGTAC CTGGATCTTG 3000  
 GGTCTTTCAC TCTAGGAGCC CAGACAAGTG GATCTGCTTG CCAGAGTCC TTTTGCCCTC 3060  
 CCTGCCACC TCCCGTGTG TCCAAGTCAG CTTTCTGCA AGAAGAAATC CTGTTAAAA 3120  
 AAGTCTTTTG TATTGGGTCA GAGTTGAAT TTGGGTGGG AGGATGGATG CAACTGAAGC 3180  
 AGAGTGTGGG TGCCAGATG TGCGCTATTA GATGTTTCTC TGATAATGTC CCCAATCATA 3240  
 CCAGGGAGAC TGGCATTGAC GAGAATCTAG GTGGAGGCTT GAGAAGGCCG AAAGGGCCCC 3300  
 TGACCTGCTC GCTTCTCTTA GCTTGCCCTT CAGCTTTGCA AAGAGCCACC CTAGGCCCCA 3360  
 GCTGACCGCA TGGGTGTGAG CCAGCTTGAG AACACTAACT ACTCAATAAA AGCGAAGGTG 3420  
 GACCNAAAAA AAAAAAAAAA AAAA

Seq ID NO: 69 Protein sequence:  
Protein Accession #: NP\_068772.1

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      MKASPRRLPI LKRRRLPLPV QNAPSETSEE EPKRSPAQQE SNQASASKEV AESNSCKFFA 60
      GIKIINHPTM PNTQVVAIPN NANIHSIITA LTAKGKESGS SGFNKFILIS CGGAPTQPPG 120
      LRPQTQTSYD AKRTEVTLET LGPKPAARDV NLFPPPGALC EQKRETADG EAAGCTINNS 180
10     LSNIQWLKRM SSDGLGSRSI KQEMEKEKNC HLEQRQVKVE EPSRPSASWQ NSVSERPPYS 240
      YMAMIQFAIN STERKRMTLK DIYTWIEDHP PYFKHIAKPG WKNSIRHNLS LHDMFVRETS 300
      ANGKVSFWTI HPSANRYLTL DQVFKPLDPG SPQLPEHLES QKRFNPPELR RNMTIKTELP 360
      LGARRKMKPL LPRVSSYLVP IQPFVNQSLV LQPSVKVPLP LAASLMSEEL ARHSKRVRIA 420
      PKVLLAEBGI APLSSAGPGK EEKLLFGEFP SPLLPVQTIK EEEIQPGEEM PHLARPIKVE 480
15     SPFLBIEWSP APSFKBESSH SWEDSSQSPT PRPKSYSGL RSPTRCVSEM LVIQHRERRE 540
      RSRSRKQHL LPPCVDEPEL LPSEGPSTSR WAAELPPPAD SSDPASQLSY SOEVGPGPKT 600
      PIKETLPISS TPSKSVLPRT PESWRLTPPA KVGGLDFSPV QTSQGASDPL FDFLGLMDLS 660
      TPLQSAAPPL ESPQLRLSSE PLDLISVPFG NSSPSDIDVP KPGSPPEQVS GLAANRSLTE 720
20     GLVLDTMDS LSKILLDISF PGLDEDPLGP DNINWSQFIP ELQ

```

Seq ID NO: 70 DNA sequence  
Nucleic Acid Accession #: BC006529.1  
Coding sequence: 178-2424

```

25     1      11      21      31      41      51
      |      |      |      |      |      |
      GGCACGAGGG GGACCCGGCC GGTCCGGGCG GAGCCCCCGT CGGGGGCCCT GGCTCGGCCC 60
30     CAGAGTGGGA GGAGCCCGGA GCCCGCCTTC GGAGCTACGG CCTAACGGGG GGGCGGACTG 120
      CAGTCTGAGG GGTCCACACT TGTGATTCTC AATGAGAGT GAAAACGCAG ATTCTAATAG 180
      AAAACTAGCC CCGCTCGGCC ACTGATTCTC AAAAGACGGA GGCTGCCCTT TCTGTTCAT 240
      AATGCCCAAA GTGAAACATC AGAGGAGGAA CCTAAGAGAT CCCCTGCCCA ACAGGAGTCT 300
      AATCAAGCAG AGGCCCTCCA GGAAGTGCA GAGTCCAAC CTGCAAGTT TCCAGCTGGG 360
35     ATCAAGATTA TTAACCAACC CACCATGCCC AACACGCAAG TAGTGGCCAT CCCCAACAAT 420
      GCTAATATTC ACAGCATCAT CACAGCACTG ACTGCCAAG GAAAAGAGAG TGGCAGTAGT 480
      GGGCCCAACA AATTCTATCT CATCAGCTGT GGGGGAGGCC CAACTCAGCT TCCAGGACTC 540
      CGGCCTCAAA CCCAAACGAG CTATGATGCC AAAAGGACAG AAGTGACCCT GGAGACCTTG 600
      GGACCAAAAC CTGACGCTAG GATGTGAAT CTTCCTAGAC CACCTGGAGC CCTTTGCGAG 660
40     CAGAAACGGG AGACCTGTGC AGATGGTGAG GCAGCAGGCT GCACTATCAA CAATAGCCTA 720
      TCCAACATCC AGTGGCTTCG AAAGATGAGT TCTGATGGAC TGGGCTCCCG CAGCATCAAG 780
      CAAGAGATGG ACGAAAAGGA GAATTGTAC CTGGAGCAGC GACAGGTAA GGTGAGGAG 840
      CCTTCGAGAC CATCAGCGTC CTGGCAGAAC TCTGTGCTG AGCGGCCACC CTACTCTTAC 900
      ATGGCCATGA TACAATTGCG CATCAACAGC ACTGAGAGGA AGCGCATGAC TTTGAAAGAC 960
45     ATCTATACGT GGATTGAGGA CCACTTTCCC TACTTTAAGC ACATTGCCAA GCCAGGCTGG 1020
      AAGAACTCCA TCQGCCACAA CCTTTCCCTG CACGACATGT TTGTCCGGGA GAGCTCTGCC 1080
      AATGGCAAGG TCTCTCTCTG GACCATTCAC CCCAGTGCCA ACCGCTACTT GACATTGGAC 1140
      CAGGTGTTTA CAGCCAGGAA ACAGCCGAAT CCAGAGCTCC GCCCGAATC GACCATCAAA 1200
      ACCGAATCCC CCTGGGCGC ACGGGGGAAG ATGAAGCCAC TGCTACCAAG GGTGAGTCA 1260
50     TACCTGGTAC CTATCCAGTT CCGCGTGAAC CAGTCACTGG TGTGCGAGCC CTCGGTGAAG 1320
      GTGCCATTGC CCTTGGGCGC TTCCCTCATG AGCTCAGAGC TTGCCCGCCA TAGCAAGCGA 1380
      FTCCGCATTG CCCCAGGGT GCTGCTAGCT GAGGAGGGGA TAGCTCCTCT TTCTTCTGCA 1440
      GGACCAAGGA AAGAGGAGAA ACTCTGTTTT GGAGAAGGGT TTCTCTCTTT GCTTCCAGTT 1500
      CAGACTATCA AGAGGAGGAA AATCCAGCCT GGGGAGGAAA TGCCCACTT AGCGAGACCC 1560
55     ATCAAAGTGG AGAGCCCTCC CTGGAAGAG TGGCCCTCCC CGGCCCATC TTTCAAAGAG 1620
      GAATCATCTC ACTCTGGGA GATTCTGTC CAATCTCCCA CCCCAAGACC CAAGAAGTCC 1680
      TACAGTGGGC TCTGCTGCCC AACCCGCTGT GTCTGGGAAA TGCTTGTGAT TCAACACAGG 1740
      GAGAGGAGGG AGAGAGCGCG GTCTCGGAGG AAACAGCATC TACTGCCTCC CTGTGTGGAT 1800
      GAGCCGGAGC TGCTCTTCTC AGAGGGGCCC AGTACTTCCC GCTGGGCGCG AGAGCTCCCG 1860
      TTCCAGCAG ACTCTCTGA CCTGCCTCC CAGCTCAGCT ACTCCAGGA AGTGGGAGGA 1920
60     CCTTTTAAAG CACCCATTAA GGAACGCTG CCCATCTCCT CCACCCCGAG CAAATCTGTC 1980
      CTCCCAGAA CCCCCTGAAT CTGGAGGCTC ACGCCCCCAG CCAAAGTAGG GGGACTGGAT 2040
      TTCAGCCAG TACAAACCCC CAGGGGTGCC TGTGACCCCT TGCCCTGACCC CTTGGGGCTG 2100
      ATGATCTCA GCAACACTCC CTGCAAAAGT GCTCCCCCCC TTGAATCACC GCAAAGGCTC 2160
      CTCAGTTTCA AACCCCTAGA CCTCATCTCC GTCCCCCTTG GCAACTCTTC TCCCTCAGAT 2220
65     ATAGAAGTCC CCAAGCCAGG CTCGCCGAG CCACAGGTTT CTGGCCTTGC AGCCAATCGT 2280
      TCTCTGACAG AAGGCCTGGT CCTGGACACA ATGAATGACA GCCTCAGCAA GATCCTGCTG 2340
      GACATCAGCT TTCTGGGCTT GGACGAGGAC CCACTGGGCC CTGACAACAT CAATCTGTCC 2400
      CAGTTTATTC CTGAGCTACA GTAGAGCCCT GCCCTTGCCC CTGTGCTCAA GCTGTCCACC 2460
70     ATCCCGGGCA CTCAAAGGCT CAGTGCAACC CAAGCCTCTG AGTGAGGACA GCAGGCGAGG 2520
      ACTGTTCTGC TCCCTATAGC TCCCTGCTGC CTGATTATGC AAAAGTAGCA GTACACCCCT 2580
      AGCACTGCT GGGACCTTGT GTTCCCAAG AGTATCTGAT TCCTCTGCTG TCCCTGCCAG 2640
      GAGCTGAAGG GTGGGAACAA CAAAGGCAAT GGTGAAAAGA GATTAGGAAC CCCCAGCCT 2700
      GTTTCATTTC TCTGCCAGC AGTCTCTTAC CTTCCCTGAT CTTTGCAGGG TGGTCCGTGT 2760
75     AAATAGTATA AATTCTCAA ATTATCTCTT AATTATAAAT GTAAGCTTAT TTCTTAGAT 2820
      CATTTATCCAG AGACTGCCAG AAGGTGGGTA GGATGACCTG GGGTTTCAAT TGACTTCTGT 2880
      TCCTTGCTTT TAGTTTGTAT AGAAGGGAAG ACCTGCAGTG CACGGTTTCT TCCAGGCTGA 2940
      GGTACCTGGA TCTTGGGTTT TCACTGCAG GGACCCAGAC AAGTGGATCT GCTTGCCAGA 3000
      GTCCCTTTTG CCCCCTCCCT CCACCTCCCT GTGTTTCCAA CTCAGCTTTC CTGCAAGAG 3060
80     AAATCCTGGT TAAAAAGTC TTTTGTATTG GGTGAGGAGT TGAATTTGGG GTGGGAGGAT 3120
      GGATGCAACT GAAGCAGAGT GTGGGTGCCG AGATGTGCGC TATTAGATGT TTCTCTGATA 3180
      ATGTCCCAAA TCATACCAGG GAGACTGGCA TTGACGAGAA CTCAGGTGGA GGTCTGAGAA 3240
      GGCGGAAAGG GCCCCGACC TGCCTGGCTT CCTTAGCTTG CCCCTCAGCT TTGCAAGAG 3300
      CCACCCTAGG CCCCAGCTGA CCGCATGGGT GTGAGCCAGC TTGAGAACAC TAACACTCTA 3360
85     ATAAAGCGA AGGTGGAAAA AAAAAAATA AAAAAA

```

Seq ID NO: 71 Protein sequence:  
Protein Accession #: AAH06529.1

	1	11	21	31	41	51	
5	MKTSRRRLI	LKRRRLPLPV	QNAPSETSEE	EPKRSPAQQE	SNQAEASKEV	AESNSCKFPA	60
	GKIINHPTM	PNTQVVAIPN	NANIHSIITA	LTAKGKESGS	SGPNKFILIS	CGGAPTQPPG	120
	LRPQTQSYD	AKRTEVTLET	LGPKAARDV	NLPRPPGALC	EQRKRETCADG	EAAGCTINNS	180
	LSNIQWLKRM	SSDGLGSRSI	KQEMEEKENC	HLEQRQVKVE	EPSRPSASWQ	NSVGERPPYS	240
	YMAMIQFAIN	STERKRMFLK	DIYTWIEDHF	PYFKHIAKPG	WKNISIRHNLS	LHDMFVRETS	300
10	ANGKVSFWTI	HPSANRYLTL	DQVFKQKRP	NPELRRNMTI	KTELPLGARR	KMKPLLPRVS	360
	SYLVPIQFPV	NQSLVLQPSV	KVFLPLAASL	MSSELARHSK	RVRIAPKVL	AEBGLAPLSS	420
	AGPGKEKKLL	FGEGFSPLLP	VQTIKEEIQ	PGEEMPHLAR	PIKVESPPLE	EWPSAPSPFK	480
	KESSHWSWEDS	SQSPTPRPKK	SYSGLRSPTR	CVSEMLVIQH	RERRERSRSR	RQHLPLPCV	540
	DEPELLFSEG	PSTERWAAEL	PPFADSSDPA	SQLSYSQEVG	GPFKTPIKET	LPISSTPSKS	600
15	VLPRTPEWSR	LTTPAKVGGL	DPSFVQTPQG	ASDPLDPDPLG	LMDLSTTPLQ	SAPPLESPQR	660
	LLSSEPLDLI	SVFPFNSSPS	DIDVFKPGSP	EPQVSGLAAN	RSLTEGLVLD	TMNDSLSKIL	720
	LDISFPGLDE	DPLGPDNINW	SQFIPELQ				

Seq ID NO: 72 DNA sequence  
Nucleic Acid Accession #: U74612.1  
Coding sequence: 178-2583

	1	11	21	31	41	51	
25	GGCAAGAGGG	GGACCCGGCC	GGTCCGGGCG	GAGCCCCCGT	COGGGGCCCT	GGCTCGGCCC	60
	CCAGGTGGGA	GGAGCCCGGA	GCCCGCCTTC	GGAGCTACGG	CCTAACGGCG	GGGGCGACTG	120
	CAGTCTGGAG	GGTCACCACT	TGTGATTCTC	AATGGAGAGT	GAAAACGCAG	ATTCTAATAG	180
	AAAACATAGC	CCCGTCGGCC	ACTGATTCTC	AAAAGACGGA	GGCTGCCCCC	TCCTGTTCAG	240
30	AATGCCCCAA	GTGAAACATC	AGAGGAGGAA	CCTAAGAGAT	CCCCTGCCCC	ACAGGAGTCT	300
	AATCAAGCAG	AGACCTCCAA	GGAAGTGGCA	GAGTCCAACT	CTTGCAAGTT	TCCAGCTGGG	360
	ATCAAGATTA	TTAAGCTGCC	CACCATGGCC	AACACGCAAG	TAGTGGCCAT	CCCCAACAA	420
	GCTAATATTC	ACAGCATCAT	CACAGCACTG	ACTGCCAAGG	GAAAAGAGAG	TGGCAGTAGT	480
	GGGCCCAACA	AATTATCTCT	CATCAGCTGT	GGGGGAGCCC	CAACTCAGCC	TCCAGGACTC	540
35	CGGCCTCAAA	CGCAAAACAG	CTATGATGCC	AAAAGGACAG	AAGTGACCTT	GGAGACCTTG	600
	GGACCAAAAC	CTGCAGCTAG	GGATGTGAAT	CTTCTAGAGC	CACCTGGAGC	CCTTTGCGAG	660
	CAGAAACGGG	AGACCTGTGC	AGATGGTGAG	GCAGCAGGCT	GCACTATCAA	CAATAGCCTA	720
	TCCAACATCC	AGTGGCTTCG	AAAGATGAGT	TCTGATGGAG	TGGGCTCCCG	CAGCATCAAG	780
	CAAGAGATGG	AGGAAAAGGA	GAATTGTCTC	CTGGAGCAGC	GACAGGTAA	GGTTGAGGAG	840
40	CCTTCGAGAC	CATCAGCGTC	CTGGCAGAAC	TCTGTGCTGT	AGCGGCCACC	CTACTCTTAC	900
	ATGCGCATGA	TACAAATCGC	CATCAACAGC	ACTGAGAGGA	AGCGCATGAC	TTTGAAGAGC	960
	ATCTATAGCT	GGATTGAGGA	CCACTTTCCC	TACTTTAAGC	ACATTGCCAA	GCCAGGCTGG	1020
	AAGAACTCCA	TCCGCCACAA	CCTTTCCCTG	CACGACATGT	TTGTCCGGGA	GACGCTGTCC	1080
	AATGGCAAGG	TCTCCTCTCT	GACCATTCAC	CCCACTGCCA	ACCGCTACTT	GACATTGGAC	1140
45	CAGGTGTGTTA	AGCCACTGGA	CCCAGGCTCT	CCACAATTGC	CCGAGCACTT	GGAATCACAG	1200
	CAGAAACGAG	CGAATCCAGA	GCTCCGCGGG	AACATGACCA	TCAAAACCGA	ACTCCCCCTG	1260
	GGCGCACGGC	GGAAGATGAA	GCCACTGCTA	CCAAGGGTCA	GCTCATACTT	GGTACCTATC	1320
	CAGTTCGCCG	TGAACCACTG	ACTGGTGTGG	CAGCCCTCCG	TGAAGGTGCC	ATTGCCCTCG	1380
	GGGGCTTCCC	TGATGAGCTC	AGAGCTTGCC	CGCCATAGCA	AGCGAGTCCG	CATTGCCCCC	1440
50	AAGGTTTTTG	GGGAACAGGT	GGTGTGTTGG	TACATGAGTA	AGTTCCTTAG	TGGCGATCTG	1500
	CGAGATTTTG	GTACACCCAT	CACCAGCTTG	TTTAAATTTA	TCTTCTTTTG	TTTATCAGTG	1560
	CTGCTAGCTG	AGGAGGGGAT	AGCTCCTCTT	TCTTCTGACG	GACCAGGGAA	AGAGGAGAAA	1620
	CTCCTGTGTT	GAGAAGGGTT	TTCTCCTTTG	CTTCCAGTTC	AGACTATCAA	GGAGGAAGAA	1680
	ATCCAGCCTG	GGGAGGAAAT	GCCACACTTA	GCGAGACCCA	TCAAAGTGGG	GAGCCCTCCC	1740
	TGGGAAGAGT	GGCCCTCCCC	GGCCCATCTT	TTCAAAGAGG	AATCATCTCA	CTCTGGGAG	1800
55	GATTGCTCCC	AATCTCCACC	CCCCAGACCC	AAGAAGTCCT	ACAGTGGGCT	TAGGTCCCCA	1860
	ACCCGGTGTG	TCTCGGAAT	GCTTGTGATT	CAACACAGGG	AGAGGAGGGA	GAGGAGCCGG	1920
	TCTCGGAGGA	AACAGCATCT	ACTGCTCCCT	TGTGTGATGT	AGCCGGAGCT	GCTCTCTCA	1980
	GAGGGGCCCA	GTACTTCCCG	CTGGGCGGCA	GAGCTCCCGT	TCCCAGCAGA	CTCCTCTGAC	2040
	CCTGCTCTCC	AGCTCAGCTA	CTCCAGGAA	GTGGGAGGAC	CTTTTAAGAC	ACCCATTAG	2100
60	GAAACGCTGC	CCATCTCCTC	CACCCCGAGC	AAATCTGTCC	TCCCAGAAAC	CCCTGAATCC	2160
	TGGAGGCTCA	CGCCCCCAGC	CAAAGTAGGG	GGACTGGATT	TCAGCCCACT	ACAAACCTCC	2220
	CAGGGTGCTC	CTGACCCCTT	GCCTGACCCC	CTGGGGCTGA	TGGATCTCAG	CACCACTCCC	2280
	TTGCAAAAGT	CTCCCCCTCC	TGAATCACCG	CAAAGGCTCC	TCAGTTTACA	ACCCTTAGAC	2340
	CTCATCTCCG	TGCCCTTTGG	CAACTCTTCT	CCCTCAGATA	TAGACGTCCC	CAAGCCAGGC	2400
65	TCCCCGAGC	CACAGTTTTC	TGGCCTTGCA	GCCATCGTTT	CTCTGACAGA	AGGCTGCTGC	2460
	CTGGACACAA	TGAATGACAG	CCTCAGCAAG	ATCCTGCTGG	ACATCAGCTT	TCTGCGCTGG	2520
	GACGAGGACC	CAGTGGGCCC	TGACAACATC	AACTGGTCCC	AGTTTATTCC	TGAGTACAG	2580
	TAGAGCCCTG	CCCTTGCCCC	TGTGCTCAAG	CTGTCCACCA	TCCCGGGCAC	TCCAAGGCTC	2640
	AGTGACCCCC	AAGCCTCTGA	GTGAGGACAG	CAGGCAGGGA	CTGTTCTGCT	CCTCATAGCT	2700
70	CCCTGCTGCG	TGATTTATGCA	AAAGTAGCAG	TCACACCCTA	GCCACTGCTG	GGACCTTGTG	2760
	TTCCCCAAGA	GTATCTGATT	CCTCTGCTGT	CCCTGCCAGG	AGCTGAAGGG	TGGGAACAAC	2820
	AAAGGCAATG	GTGAAAAGAG	ATTAGGAACC	CCCCAGCCTG	TTTCCATTCT	CTGCCAGCA	2880
	GTCTCTTACC	TTCCTGAGT	TTTGAGGGGT	GGTCCGTGTA	AATAGTATAA	ATTCTCCAAA	2940
	TTATCTCTTA	ATTATAAATG	TAAGCTTAAT	TCCTTAGATC	ATTATCCAGA	GACTGCCAGA	3000
75	AGGTGGGTAG	GATGACCTGG	GGTTTCAATT	GACTTCTGTT	CCTTGTCTTT	AGTTTGTATA	3060
	GAAGGGAAGA	CCTGCAGTGC	ACGGTTTCTT	CCAGGCTGAG	GTACCTGGAT	CTTGGGTCTT	3120
	TCATCTCAGG	GACCCAGACA	AGTGGATCTG	CTTGCCAGAG	TCCTTTTTCG	CCCTCCCTGC	3180
	CACCTTCCCG	TGTTTCCAAG	TCAGCTTTCC	TGCAAGAGA	AATCCTGGTT	AAAAAGTCT	3240
	TTTGTATTGG	GTACAGGATT	GAATTGGGGG	TGGGAGGATG	GATGCAACTG	AAGCAGAGTG	3300
80	TGGGTGCCCA	GATGTGCGCT	ATTAGATGTT	TCTCTGATAA	TGTCGCCAAT	CATACCAGGG	3360
	AGACTGCGAT	TGACGAGAAC	TCAGGTGGAG	GCTTGAGAAG	GCCGAAAGGG	CCCTGACCT	3420
	GCCTGGCTTC	CTTAGCTTGC	CCCTCAGCTT	TGCAAGAGGC	CACCTAGGC	CCAGCTGAC	3480
	CGCATGGGTG	TGAGCCAGCT	TGAGAACACT	AACTACTCAA	TAAAGCGAA	GGTGACAAA	3540
85	AAAAAAAAAA	AAAAA					

Seq ID NO: 73 Protein sequence:  
Protein Accession #: AAC51128.1



1 11 21 31 41 51  
 5 MKTSRRRLPI LKRRRLPLPV QNAPSETSEE EPKRSPAQQE SNQAEASKEV AESNSCKFFPA 60  
 GIKIINHEPTM PNTQVVAIPN NANIHSIITA LTAKGKESGS SGPKNKFILIS CGGAPTQPPG 120  
 LRPTQTSTSYD AKRTEVILET LGPKPAARDV NLPRPPGALC EQKRETCADG EAAGCTINNS 180  
 LSNIQWLKRM SSDGLGSRSI KQEMKEEKENC HLEQRQVKVE EPSRPSASWQ NSVSERPPYS 240  
 YMAIQAFAIN STERKMTLK DIYTWIEDHF PYFKHIAKPG WKNSIRHNL LHMDFVRETS 300  
 10 ANGKVSFWTI HPSANRYLT DQVFKPLDPG SPQLPEHLES QOKRPNPELR RNMTIKTELP 360  
 LGARRKMKPL LERVSSYLVP IQPPVQSLV LQPSVKVPLP LAASLMSEL ARHSKRVRIA 420  
 PKVFGQVVF GYMSKFFSGD LRDPGTFTS LFNFIPLCLS VLLAEGLAP LSSAGPGKEE 480  
 KLLFGEGFSP LLPVQTIKEE EIQPGEMPH LARPIKVESF PLEWPSAP SPKEESSHW 540  
 EDSSQSPTR PKKSYSLRS PTRCVSEMLV IQHRRRERS RSRKQHLPL PCVDEPELLF 600  
 15 SEGSPSTRWA AELFFPADSS DPASQLSYSQ EVGGPFKTP KETLPISTP SKSVLFRTPF 660  
 SWRLTPPAKV GGLDFSVQT SQGASDPLPD PLGLMDLSTT PLQSAFFLES PQRLLSSEPL 720  
 DLISVFPFNS SPSSIDVFKP GSPEPQVSL AANRSLTEGL VLDTMNDSL KILLDISFPF 780  
 LDEDFLPGDN INWSQFIPEL Q

Seq ID NO: 74 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 111-416

1 11 21 31 41 51  
 25 GGGAAGAGCC AGGCTGAGCC TTATAAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60  
 TCATCCTTCT ACTCGTGAGC CTTCAGCT CTGCTTTTTT GAAAGCAAAG ATGAGCAACA 120  
 CTCAGCTGA GAGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGTG 180  
 ATGACAGAT TGAGAAGCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAACTTCC 240  
 30 TTAGTGCTCT TGACAAAAG GGCACAAAT ACCTGCGCA TGTCTTTGAG AAAAAGGACA 300  
 AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCA 360  
 CAGACTACCA CAGCAGAGC CATGAGCAG CGCCTGTTC CGGGGCGAGC CAGTGACCCA 420  
 GCCCCACCAA TGGCCTCCA GAGACCCAG GAACAATAA ATGTCTTCTC CCACCAGA

Seq ID NO: 75 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 40 MSNTQAERSI IGMIDMFHYK TRDDKIEKP SLLTMMKENF PNFLSACDKK GTNYLADVFE 60  
 KKDKNEDKKI DFSEFLSLIG DIATDYHKQS HGAAPCSGGS Q

Seq ID NO: 76 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 111-416

1 11 21 31 41 51  
 50 GGGAAGAGCC AGGCTGAGCC TTATAAAGGA CTGCTCTTTG TCCAAACACA CACATCTCAC 60  
 TCATCCTTCT ACTCGTGACA CTTCAGCT CTGCTTTTTT GAAAGCAAAG ATGAGCAACA 120  
 CTCAGCTGA GAGTCCATA ATAGGCATGA TCGACATGTT TCACAAATAC ACCGACGTG 180  
 ATGCAAGAT TGAGAAGCCA AGCCTGCTGA CGATGATGAA GGAGAACTTC CCCAACTTCC 240  
 TCAGTGCTCT TGACAAAAG GGCATACATT ACCTGCGCAC TGTCTTTGAG AAAAAGGACA 300  
 55 AGAATGAGGA TAAGAAGATT GATTTTCTG AGTTTCTGTC CTGCTGGGA GACATAGCCG 360  
 CAGACTACCA CAGCAGAGC CATGAGCGG CGCCTGTTC TGGGGGAAGC CAGTGATCCA 420  
 GCCCCACCAA GGGCCTCCA GAGACCCAG GAACAATAA TGTCTCTCC CACCAGA

Seq ID NO: 77 Protein sequence:  
 Protein Accession #: XP\_048124.1

1 11 21 31 41 51  
 60 MSNTQAERSI IGMIDMFHYK TGRDGKIEKP SLLTMMKENF PNFLSACDKK GIHYLATVFE 60  
 65 KKDKNEDKKI DFSEFLSLIG DIAADYHKQS HGAAPCSGGS Q

Seq ID NO: 78 DNA sequence  
 Nucleic Acid Accession #: Z73678.1  
 Coding sequence: 253-2433

1 11 21 31 41 51  
 75 GGGGTGGTGC AGGGCAGGGG TGGTATATCC TGTCTGACGG AGGGCGGGCC TGGCCAGTGC 60  
 CAGAGAGGGA CGAACCAGGG TGGAGCGGCC AGGAGCAGCT GCAGGGAGCC CTCACGCGGA 120  
 CCTCGCACTC TATGGCCGTA GGGAGCGGCT GAGAGCGAGA AGAGCAAGCT CCTGCCCGCC 180  
 CGCTGCACCG CACTCGCCT CGCCTCTCTG CTCTCCTAGG CCGCGGCGC GCGCCACCGC 240  
 CCTCCGCGCA CATGAACCA CTGCGGCTC AAGACCGCT TGGCGTACA ATGCTTCCAG 300  
 GACCAGGACA ACTCCACGTT GGCTTTGCCG TCGGACCAA AGATGAAAC AGGCACGTCT 360  
 80 GGCAGGCAGC GCGTGACAGA GCAGGTGATG ATGACCGTCA AGCGGCAGAA GTTCAAGTCT 420  
 TCCAGTCGT CCACCTGAG CCACTCCAAT CGAGGTTCCA TGTATGATGG CTTGGCTGAC 480  
 AATTACAAT ATGGGACAC CAGCAGGAGC AGCTACTACT CCAAGTTCCA GGCAGGGAAT 540  
 GGCTCATGGG GATATCCGAT CTACAATGGA ACCCTCAAGC GGGAGCCTGA CAACAGGCGC 600  
 TTCACTCTCT ACAGCCAGAT GGAGAACTGG AGCGGCACT ACCCCCGGGG CAGCTGTAA 660  
 85 ACCACCGCGC CAGGCAGCA CATCTGCTTC ATGCAGAAA TCAAGGGGAG CGCAGTGAG 720  
 CCGACCTCT ACTGTGACCC ACGGGGCAAC CTGCGCAAGG GCAAGCTGGG CAGCAAGGGC 780  
 CAGAAGACCA CCGAAGACCG CTACAGCTTT TACAGCACT GCAAGTGTCA GAAGGCCATA 840  
 AAGAAGTGCC CTGTGCGGCC GCCCTCTGT GCTTCCAAGC AGGACCTGT GTATATCCC 900



	CCCATCTCCT	GCAACAGGA	CCTGTCTTT	GGCCACTCTA	GGCCAGCTC	CAAGATCTGC	960
	AGTGAGGACA	TGAGGTGAC	TGGGCTGACC	ATCCCCAAGG	CTGTGACGTA	CCTGAGCTCC	1020
	CAGGATGAGA	AGTACACGGC	CATTGGGGCC	TATTACATCC	AGCATACCTG	CTTCCAGGAT	1080
5	GAATCTGCGA	AGCAACAGGT	CTATCAGCTG	GGAGGCATCT	GCAAGCTGGT	GGACCTCCTC	1140
	CGCAGCCCCA	ACCAGAACGT	CCAGCAGGCC	GGCGCAGGGG	CCCTGCGCAA	CCTGGTGTTC	1200
	AGGAGCACCA	CCAAACAGCT	GGAGACCCGG	AGGCAGAAATG	GGATCCGGGA	GGCAGTCAGC	1260
	CTCCTGAGGA	GAACCGGGAA	CGCCGAGATC	CAGAAGCAGC	TGACTGGGCT	GCTCTGGAAC	1320
	CTGTCTTCCA	CTGAGAGCT	GAAGGAGGAA	CTCATTGCGG	ACGCCCTGCC	TGTTCTGGCC	1380
10	GACCGCGTCA	TCAITCCCCT	CTCTGGCTGG	TGCGATGGCA	ATAGCAACAT	GTCCCGGGAA	1440
	GTGGTGACC	CTGAGGTCTT	CTTCAATGCC	ACAGGCTGCT	TGAGGAACCT	GAGCTCGGCC	1500
	GATGAGGCC	GCCAGACCAT	GGTAACTAC	TCAGGGCTCA	TTGATTCCCT	CATGGCCTAT	1560
	GTCCAGAACT	GTGTAGCGGC	CAGCCGCTGT	GACGACAAST	CTGTGAAAAA	CTGCATGTGT	1620
	TTCTTGACCA	ACCTCTCCTA	CCGCCCTGAC	GCCGAGGTGC	CCACCCGCTA	CGCCAGCTG	1680
15	GAGTATAACG	CCCGCAACG	CTACACCGAG	AAGTCCCTCA	CTGGCTGCTT	CAGCAACAA	1740
	AGGCAACAGA	TGATGAACAA	CAACTATGAC	TGCCCCCTGC	CTGAGGAAGA	GACCAACCCC	1800
	AAGGGCAGCG	GCTGGTTGTA	CCATTGAGAT	GCCATCCGCA	CCTACCTGAA	CCTCATGGGC	1860
	AAGAGCAAGA	AAGATGCTAC	CTCTGAGGCC	TGTGCTGGTG	CCCTGAGAAA	CCTGACAGCC	1920
	AGCAAGGGTG	TGATGTCCAG	TGGCATGAGC	CAGTGTGATTG	GGCTGAAGGA	AAAGGGCCTG	1980
20	CCACAAATTG	CCGCCCTCCT	GCAATCTGGC	AACCTCTGATG	TGGTGGCTGT	CGGAGCCTCC	2040
	CTCCTGAGCA	ACATGTCCCG	CCACCCTCTG	CTGCACAGAG	TGATGGGGAA	CCAGGTGTTC	2100
	CCGAGGTGTA	CCAGGCTCCT	CACGAGCCAC	ACTGGCAATA	CCAGCAACTC	CGAAGACATC	2160
	TTGTCTCTGG	CCTGCTACAC	TGTGAGGAAC	CTGATGGCCT	CGCAGCCACA	ACTGGCTAAG	2220
	CAGTACTTCT	CCAGCAGCAT	GCTCAACAAC	ATCATCAACC	TGTGCGGAAG	CAGTGCTCTA	2280
25	CCCAAGGCTG	CAGAAGCTGC	CCGGCTTCTC	CTGTCTGACA	TGTGCTCAG	CAGGAACCTG	2340
	CAGGGTGTC	TCAGACAGCA	AGGTTTCGAT	AGGAACATGC	TGGGAACCTT	AGCTGGGGCC	2400
	AACAGCTTCA	GAGACTTCA	CTCCGATTTC	TAAGAAGAGA	CTGTCCAAAG	AAGTTAGGCT	2460
	TGCAAGAAAG	TATGACCAG	CTGAGAACCC	CTCAGGCCTC	GCTGGATGGG	GTCTTCTGTC	2520
	CATCTCTGTC	AGTATTGGG	AAAGTTCACA	AGAAACTGAG	AAGAACTTA	AAAACCTGGG	2580
30	ATAGTGAAAG	GATTTTATGA	TTTTTTTTTT	CCTTGGGGAA	ACTGGCAGGC	AATGGGGGTT	2640
	AGGAGGTGTC	GGGCGGGGGG	GGCTTTCTTG	AGTTAAAGGG	GCTTATATGT	GATGTCAATA	2700
	TTTCTTCTCT	TGAGAAATGG	TATATATATG	TGTCTAATGT	AAGTGTGTGC	ATGCATGTGC	2760
	GGGTGATGT	GTGTGTGTGT	GAGTGTCTTA	AAGCATAAAC	ACAACTGCA	AAAAGCTAGG	2820
	TAAGCTATT	TGTTGACGCT	CATAAGGTGG	TGAAAAGGAC	TCTCTGTGT	TTCTTACTCA	2880
35	TAGGCAAGGA	CAACATGTGC	TTTTTGGTGA	GCTGCTCATA	ATTCTGAAA	TGTGTGTGTC	2940
	CAGGSCAAGS	GGGCGATCAC	TGCAGTCAGG	CCCTCAGAGG	AGTCTGTCAG	GCTTCTTACC	3000
	AGTGTCTTCC	AAGGCTGACG	GAGTAACTGG	GGCTGGGGCA	GCCTCCCCCC	TTACAAGGCT	3060
	GCTTTCACG	AAGGAGGCTC	TGTTGTATCT	CATGGGAGAA	TCTGGGGTGT	CTGTAGTGTG	3120
	ACCCCTCCAG	CAGCGCCACA	AGGACTGAGG	TTGGGTAGGT	GTGAGGTTCC	AGAGGACAGC	3180
40	AGGACACTCT	CGCATACTTT	GCCAAATGAG	GCTGCTCAG	AGGAGTAGGA	GCTGAAAGAT	3240
	GGTGCTTCC	AGCTCTTGGG	GCTGTGTGCC	CATCAGAGCA	GGCTCAGCCT	GCAAGGCCCC	3300
	TGCAATCAGA	GATCTTTGTA	TCTACTTGTT	GCAGGAGAAA	GAAGGTAAAA	AATGATTTTT	3360
	TTAAGAAAG	CTATTTTTAT	CGAGCTCTTT	CCCAAGAGCT	GTCTGAGGAA	TGGCTGGTCT	3420
	TCATATTCCC	AGTGAGAGGG	GGAACAAGTG	GGGCTGGGCA	TATACCTATT	CCGGCTTCTA	3480
45	GTGGATGGA	TTGGGGGTAT	AGAAATTAAC	CAGGAAGATG	TTTCCACCAA	GCCTGCTGTG	3540
	AGTCAATTGA	GGGAGTGTTT	GGGTCCCAGG	AGACTTGGAC	GGGGGGAGTT	TGGGTAGACT	3600
	AGGAAGAGAA	AGTGCCATAT	CAGGGTACCG	GTACCGGCAA	GCTCACATCT	CAGCCAGGGG	3660
	CCATGCCCCA	CTTCCCTGTA	CCCCAGCTGT	CTTGTCTCCA	CTCTGTGAAA	CCCAAGGGGG	3720
	ATGTGATAAA	CAGGGCTATT	AGGGGTATCA	GCCACGTGGA	GGCCCCAGAC	TCTGTGCACT	3780
50	TCAGACCAGC	AGCAGCAGGA	GGGCTCCGGA	GGGCCCTTATG	AGAAAACCTG	TGTGACATC	3840
	CCTTGTGTGA	CATTAAGACA	GAGCAGAGCC	CAGCGCTCCC	AAGCCTTCTT	CCTTCCAGCT	3900
	TCTACTTCCA	TGCTAGCAAT	GCTGTGTGTA	GAGAGGAATT	AACCTCTTGG	TCTGTGCCCT	3960
	CTCTAGAAAG	AAATATAAGAT	GCTCTCTCTC	CTCACCCCTT	CTCAGCCTCC	TCCCAAGTCT	4020
	TCCTCTTCTG	CACCAACCCC	GAGTCCAAAC	CCACCTCTTG	CCCCAGCATT	CAGGCTGGAA	4080
55	AACACTGATG	TGGACTCAGT	ATGACAACTG	AGATGGGGGA	AGCCAGACAT	GTGAGGACGC	4140
	TGTCCTCCGA	GAGGTGTCCC	CGGCTGTTAG	CCAGCTGTGC	TGTGTGTCTG	TGGGTCTGTC	4200
	ATACCTCTCC	TGCTTCTGTG	TCACACTGGG	AGGCCCACTC	CTGGCTCACC	TCTCCTCTCT	4260
	AGGGAACCCAC	GTGGGAGCCT	GGATCCCTGG	ACTGTCTCTG	GCATAGGTTT	CAGGGGCTCT	4320
	CTTTGTGTCT	ATCAGAACCC	AGAGGAATTG	TTCTCCTAAA	AAATACGTAT	GGCATACCAA	4380
60	TCTGTGCGGG	GCATGTCTCT	AAGCACTTAG	ACTACATCAG	GGAAGAACAC	AGACCAATC	4440
	CCCGTCCCTA	TGCGGCTTAT	GTTTTCTGGA	GGAAAGTGGG	GACACAAGTC	CTTGGCTTTA	4500
	GGGCTCCCCC	GGCTGGGGGC	TGTGAGTCC	GGTCAGGGCG	GGAGGGGAAA	TGCACCGCTG	4560
	CATGTGAACC	TTACAGACCC	AGGCGGATGC	CCCTTCCCTT	TAGCACTACC	CTGGCTCTCT	4620
	GCATCCCTCT	GCTCATGTGT	CCTCCACCTT	TCAAAGAAATG	AAGAGCCCCA	TGGGCCCAGC	4680
65	CCCTGCCCTG	GGAACCAAGC	AGCCTTCCAG	ACCTCAGGGG	CTGAGGCAGA	CTATTAGGGC	4740
	AGGGCTGACT	TTGGTGACAC	TGCCCATTTCC	CTCTCAGGCC	AGCTCAGGTC	ACCCGGGCCT	4800
	CTGACCCAGG	CCTGTCACTT	TGAGAGGGGC	AAAACCTAGA	GGGGCTTTTC	CTAGAGAAAG	4860
	AGAACAAGGA	GCTTGCCAGG	CTTCAATGTAG	CCGACACAAG	TCTCAGGATT	TTAAGTCCAC	4920
	ATTGGCTTCA	CATAGCCCTA	GGCCAAATGCC	CAAAATAAGG	AGTTCCAATT	TGGGGCCAAA	4980
70	TGAGGAAGGA	CACAGACTCT	GCCCTGGGAT	CTCCTGTGCT	AGCGGCCAAT	GACAAATCCA	5040
	GTCATTGGCC	ACCAGCCACC	TCTGAGTGG	GGACCAACT	AGCAGCCCTG	ACTCCCACT	5100
	CCTCTGGGG	ACCCAAGAGG	CAGTGTGTCT	GTCTGCGTGT	CCACCTTGGG	ATCTGGCTGA	5160
	ACTGGCTGGG	AGGACCAAGA	CTGCGGCTGG	GGTGGGCAGG	GAAGGGAAGC	CGGGGGCTGC	5220
	TGTGAGGGAT	CTTGGAGCTT	CCCTGTAGCC	CACCTTCCCT	TTGCTTCATG	TTTGTAGAGG	5280
75	AACCTTGTGC	CGGCGAGGCC	CAGTTTCTCT	GTGTGATACA	CTAATGTATT	TGCTTTTTTT	5340
	GGAAATAGAG	AAATCAATA	AATTGCTAGT	GTTCCTTTGA	AAAAAATA		

Seq ID NO: 79 Protein sequence:

Protein Accession #: CAA98022.1

80	1	11	21	31	41	51	
	MNHSPLKAL	AYECFQDQDN	STLALPSDQK	MKTGTSGRQR	VQEQVMITVK	RQSKSSQSS	60
	TLHSNRRGSM	YDGLADNYNY	GTTSRSSYYS	KPQAGNGSWG	YPIYNGTLKR	EPDNRFFSSY	120
85	SQMNWSRHY	PRGSCNTTGA	GSDFCFMQKI	KASRSEPDLY	CDPRGTLRKG	TLGSKGQKTI	180
	QNRYSFYSTC	SGQKAIKKCP	VRPPSCASKQ	DPVYIPPIISC	NKDLSPFHSR	ASSKICSEDI	240
	ECSGLTIPKA	VQYLSQDEK	YQAIGAYYIQ	HTCFQDESAK	QVYQLGGIC	KLVDLLRSPN	300
	QNVQAAAGA	LRNLVFRSTT	NKLETRRQNG	IREAVSLLR	TGNABIQQL	TGLLNNLSST	360

DELKEELIAD ALFVLADRIV IPFSGWCDGN SNMSREVVDV EVFPNATGCL RNLSSADAGR 420  
 QTMRYNSGLI DSLMAYVQNC VAASRCDDKS VENCMLVHNL LSYRLDARVP TRYRQLEYNA 480  
 RNATYKSSST GCFPNKSKDM MNMNYDCPLP EETNPNGSG WLYHSDAIRT YLNLMGKSKK 540  
 5 DATEACAGA LQNLTAASKGL MSSGMSQLIG LKEKGLPQIA RLLQSGNSDV VRSGASLLSN 600  
 MSRHPLLRHV MGNQVFPFVLT RLLTSRTGNT SNSEILSSA CYTVRNLMAS QPQLAKQYFS 660  
 SSMLNINIINL CRSSASPKAA EAARLLLSDM WSSKELQGV L RQQGFDRNML GTLAGANSLR 720  
 NFTSRF

Seq ID NO: 80 DNA sequence  
 Nucleic Acid Accession #: NM\_006516.1  
 Coding sequence: 180-1658

1 11 21 31 41 51  
 TAGTCGCGGG TCCCGAGTG AGCAGCCAG GGAGCAGGAG ACCAAACGAC GGGGGTCGGA 60  
 GTCAGAGTCG CAGTGGGAGT CCCCAGACCG GAGCAGCAGC CTGAGCGGGA GAGCGCCGCT 120  
 CGCAGCGCCG TCGCCACCCG CGTACCCGCG GAGCAGCAGG CCACCAGCGC AGCGCTGCCA 180  
 20 TGGAGCCGAG CAGCAAGAAG CTGACGGGTC GCCTCATGCT GGCTGTGGGA GGAGCAGTGC 240  
 TTGGCTCCCT GCAGTTTGGC TACACACTG GAGTCATCAA TGCCCCCAGG AAGGTGATCG 300  
 AGGAGTTCTA CAACCAGACA TGGGTCCACC GCTATGGGGA GAGCATCTCG CCCACCACGC 360  
 TCACCACGCT CTGGTCCCTC TCAGTGGCCA TCTTTTCTGT TGGGGGCATG ATTGCTCTCT 420  
 TCTCTGTGGG CCTTTTGGTT AACCGCTTTG GCGGCGGAA TTCAATGCTG ATGATGAACC 480  
 25 TGCTGGCCTT CGTGTCCGCC GTGCTCATGG GCTTCTCGAA ACTGGGCAAG TCCTTTGAGA 540  
 TGTGATCTCT GGGCGGCTTC ATCATCGGTG TGTACTCGGG CCTGACCAACA GGCTTGTGTC 600  
 CCATGTATGT GGGTGAAGTG TCACCACAG CCTTTCTGTG GGGCCTGGGC ACCCTGCACC 660  
 AGCTGGGCGT CGTCTCGGCG ATCCTCATCG CCCAGGTGTT CGGCTTGGAC TCCATCATGG 720  
 GCAACAAGGA CCGTGGGCCC CTGCTGCTGA GCATCATCTT CATCCCGGCC CTGCTGCAGT 780  
 30 GCATCTGTCT CGCCTTCTGC CCGGAGAGTC CCGCTTCTCT GCTCATCAAC CGCAACGAGG 840  
 AGAACCGGGC CAAGAGTGTG CTAAGAAGAG TGCGCGGGAC AGCTGACGTG ACCCATGACC 900  
 TCAGAGGATG GAAGCAAGAG AGTCGGGAGA TGATGCGGGA GAAGAAGGTC ACCATCTCGG 960  
 AGCTGTTCCG CTCGCCCGCC TACCGCCAGC CCATCCTCAT CGCTGTGGTG CTGCAGCTGT 1020  
 CCCAGCAGCT GTCTGGCCTC AACGCTGTCT TCTATTACTC CACGAGCATC TTCGAGAAGG 1080  
 35 CGGGGGTGCA CAGCCTGTG TATGCCACCA TTGGCTCGGG TATGCTCAAC ACGGCTTCA 1140  
 CTGTCTGTCT GCTGTTTGTG GTGAGGAGAG CAGGCGGGCG GACCTTGCAC CTCATAGGCC 1200  
 TCCTGTGGAT GCGCGGTTGT GCCATCTCA TGACCATCGC CTAGCAGTGC CTGAGCAGC 1260  
 TACCTCGGAT GTCTATCTG AGCATGTGG CCATCTTGG CTTTGTGGCC TTCTTTGAAG 1320  
 TGGTCTCTGG CCCCATCCCA TGGTTCATCG TGGCTGAAC CTTCAGCCAG GGTCCACGTC 1380  
 40 CAGCTGCCAT TGCGTTGCA GGCTTCTCCA ACTGGACCTC AAATTTCAIT GTGGGCTGT 1440  
 GCTTCCAGTA TGTGAGCAA CTGTGTGGTC CCTAGTCTT CATCATCTC ACTGTGTCTC 1500  
 TGGTCTGTGT CTTCATCTTC ACCTACTTCA AAGTTCCTGA GACTAAAGGC CGGACCTTGC 1560  
 ATGAGATCGC TTCCGGCTTC CGGAGGGGG GAGCCAGCCA AAGTGATAAG ACACCCGAGG 1620  
 AGCTGTTCOA TCCCTCGGG GCTGATTCCC AAGTGTGAGT CGCCCCAGAT CACGACCCG 1680  
 45 GCCTGTCTCC AGCAGCCCTA AGGATCTCTC AGGAGCACAG GCAGCTGGAT GAGACTTCCA 1740  
 AACCTGACAG ATGTGACGCG AGCCGGGCTT GGGGCTCCTT TCTCCAGCCA GCAATGATGT 1800  
 CAGAGAAGAT ATTACAGACT TAACGGCTCC AGGATTTTAA CAAAAGCAAG ACTGTGCTC 1860  
 AAATCTATTC AGACAAGCAA CAGGTTTAT AATTTTITA TTACTGATTT TGTATTTTT 1920  
 ATATCAGCTC GAGTCTCTCT TGCCACATC CCAGGCTTCA CCTGTAATGG TTCCATGCTC 1980  
 50 GAGGGTGGAG ACTAAGCCCT GTGAGACAC TTGCTTCTT CACCCAGCTA ATCTGTAGGG 2040  
 CTGGACCTAT GTCCCTAAGGA CACACTAATC GAACATGAA CTACAAAGCT TCTATCCAG 2100  
 GAGGTGGCTA GTTGGACCG TTCTGTGGC CTGGATCTCC CCACTTAGG GGTGAGGCTC 2160  
 CATTAGGATT TGCCCTTCC CATCTCTTCC TACCCAACCA CTCAAATTA TCTTCTTTA 2220  
 CCTGAGACCA GTTGGGAGCA CTGGAGTGA GGGAGGAGG GGAAGGGCC AGTCTGGGCT 2280  
 55 GCGGGTTCT AGTCTCTTT GCACTGAGGG CCACTATT ACCATGAGAA GAGGCGCTGT 2340  
 GGGAGCCTGC AAATCTACTG CTCAGAAGA CATGGAGACT CCGCCCTGT TGTGTATAGA 2400  
 TGCAAGATAT TTATATATAT TTTTGGTTGT CAATATTAA TACAGACACT AAGTTATAGT 2460  
 ATATCTGGAC AAGCCAACT GTAAATACAC CACCTCACTC CTGTTACTTA CCTAAACAGA 2520  
 TATAAATGCG TGGTTTITAG AAACATGGTT TTGAAATGCT TGTGGATTGA GGGTAGGAGG 2580  
 60 TTTGGATGGG AGTGAGACAG AAGTAAGTGG GGTGCAACC ACTGCAACGG CTAGACTCT 2640  
 GACTCAGAT CCAGTCCCTT ACACGTACCT CTCATCAGTG TCCTCTGTCT CAAAATCTG 2700  
 TTTGATCCCT GTTACCCAGA GAATATATAC ATTCTTATC TTGACATTCA AGGCATTTCT 2760  
 ATCACATATT TGATAGTTG TGTTCAAAA AACACTAGTT TTGTGCCAGC CGTGATGCTC 2820  
 AGGCTTGAAA TGCATTATT TTGAATGTGA AGGGAA

Seq ID NO: 81 Protein sequence:  
 Protein Accession #: NP\_006507.1

1 11 21 31 41 51  
 MEPSSKLTG RIMLAVGAV LQSLQFGYNT GVINAPOKVI EEFYNQIVWH RYGESILEPT 60  
 LITLWLSLVA IPSVGGMIGS FSVGLFVNRF GRRNSMLMNM LLAPVSAVLM GFSKLKKSFE 120  
 MLILGRFIIG VYGLTTGTFV PMYGEVSPT AFRGALGTLH QLGIIVGILI AQVPLGDSIM 180  
 75 GNKDLWPLL SIIPFALLQ CIVLFPFES PRFLINRNE ENRAKSVLKK LRGTADVTHD 240  
 LQEMKEESRQ MMREKVTIL ELFRSPAYRQ PILIAVVLQ SQQLSGINAV FYYSTSIFEK 300  
 AGVQOPVYAT IGSIVNTAF TVVSLFVVER AGRTLHLIG LAGMAGCAIL MTIALALLEQ 360  
 LPWMSYLSIV AIPGFVAFPE VGPGLIPWFI VAELEFSQGR PAIAVAGFS NWTNFIIVGM 420  
 CPYVEQLCG PYVFIIFTVL LVLFFIFTYP KVPETKGRTP DEIASGFRQG GASQSDKTPE 480  
 ELFHPLGADS QV

Seq ID NO: 82 DNA sequence  
 Nucleic Acid Accession #: BC001291  
 Coding sequence: 44-541

1 11 21 31 41 51  
 GGGGGCGCCG CGCGCTGACC CTCCTGGGCG ACCGCTGGGG ACGATGGGCG TGCTCGCCTT 60  
 85 GCTGCTGGTC GTGGCCCTAC CCGGGGTGTG GACAGACGCC AACCTGACTG CGAGACAAAC 120

AGATCCAGAG GACTCCCAGC GAACGGACGA GGGTGACAA AGAGTGTGGT GTCATGTTG 180  
 TGAGAGAGAA AACACTTTTC AGTGCCAGAA CCCAAGGAGG TGCAAAATGA CAGAGCCATA 240  
 CTGCGTTTATA GCGGCCGTGA AAATATTTCC ACGTTTTTTC ATGGTTTGGG AGCAGTGCTC 300  
 CGCTGGTTGT GCAGCGATGG AGAGACCCAA GCCAGAGGAG AAGCGGTTTC TCCTGGGAAGA 360  
 GCGCATGCCC TTCTTTTACC TCAAGTGTG TAAATTTCCG TACTGCAATT TAGAGGGGGC 420  
 ACCTATCAAC TCATCAGTGT TCAAAGAATA TGCTGGGAGC ATGGGTGAGA GCTGTGGTGG 480  
 GCTGTGGCTG GCCATCTCC TGCTGCTGGC CTCCATTGCA GCGGCGCTCA GCCTGTCTTG 540  
 AGCCACGGGA CTGCCACAGA CTGAGCCTTC CGGAGCATGG ACTCGCTCCA GACCGTTGTC 600  
 ACCTGTTGCA TTAAACTTGT TTTCTGTTGA TTACCTCTTG GTTTGACTTC CCAGGGTCTT 660  
 GGGATGGGAG AGTGGGGATC AGGTGCAGTT GGCTCTTAAC CCTCAAGGGT TCTTTAACTC 720  
 ACATTCAAGG GAAGTCCAGA TCTCCTGAGT AGTGATTTTG GTGACAAGTT TTTCTCTTG 780  
 AAATCAAACC TTGTAATCA TTTATTGCTG ATGGCCACTC TTTTCTTGA CTCCTCTG 840  
 CCTCTGAGGG CTTCAGTATT GATGGGGAGG GAGGCCTAAG TACCCTCAT GGAGAGTATG 900  
 TGCTGAGATG TTCCGACCT TTCAGGTGAC GCAGGAACAC TGGGGGAGTC TGAATGATTG 960  
 GGGTGAAGAC ATCCCTGGAG TGAAGGACTC CTCAGCATGG GGGGCGAGTG GGCACACGTT 1020  
 AGGGCTGCCC CCAATTCAGT GGTGGAGGCG CTGTGATGG CTGCTTTTCC TCAACCTTTC 1080  
 CTACCAGATT CCAGGAGGCA GAAGATAACT AATTGTGTGG AAGAACTTA GACTTCAACC 1140  
 ACAGCTGGC ACAGGTGCAC AGATTCATAA ATTCACACAC GTGTGTGTTT AACATCTGAA 1200  
 ACTTAGGCCA ACTAGAGAGC ATCAGGGTAA ATGGCGTTC TTTCTCTGTT AAGATGCAGC 1260  
 CATCCATGGG GAGCTGAGAA ATCAGACTCA AAGTTCACCC AAAACCAAT ACAAGGGGAC 1320  
 TTCAAAGTT CACGAAAAA AAAAAAATA AAAAAAATA AAAAAAATA AAA

Seq ID NO: 83 Protein sequence:  
 Protein Accession #: AAH01291

1 11 21 31 41 51  
 MALLALLLVV ALPRVWTDAN LTARQDPED SQRTEGDNR VMCHVCEREN TFEQNPERRC 60  
 KWTPEYCVIA AVKIFPPFFM VAKQCSAGCA AMERPKPEEK RFLLEPMPPF FYLKCKKIRY 120  
 CNLEGPINS SVFKEYAGSM GESCGGLWLA ILLLLASTIAA GLSL

Seq ID NO: 84 DNA sequence  
 Nucleic Acid Accession #: NM\_022893.1  
 Coding sequence: 229-2726

1 11 21 31 41 51  
 TTTTITTTTT TTTTITGCTT AAAAAAAGC CATGACGGCT CTCCACAAT TCATCTTCCC 60  
 TGCGCATCT TGTATTATT TCTAATTTAT TTTGGATGTC AAAAGGCACT GATGAAGATA 120  
 TTTTCTCTGG AGTCTCTTTC TTTCTAACCC GGCTCTCCCG ATGTGAACCG AGCGGTGCTC 180  
 GCGCGCGCGC CGCGCGCGCG GCGCGCGCGC CCGCGCGCGC AGCCCAACAT GTCTCGCGCG 240  
 AAGCAAGGCA AACCCGAGCA CTTAAGCAAA CGGGAATTCT GCGCGAGCC TCTTGAAGCC 300  
 ATTCTTACAG ATGATGAACC AGACCAAGCG CCGTTGGGAG CTCACAGAAG GGATCATGAC 360  
 CTCTCACCT GTGGGCGAGT CCAGATGAAC TTCCCATTGG GGGACATTCT TATTTTATC 420  
 GAGCACAAAC GGAAACAATC CAATGGCAGC CTCTGCTTAG AAAAGCTGT GGATAAGCCA 480  
 CCTTCCCTTT CACCAATCGA GATGAATAAA GCATCCAATC CCGTGGAGGT TGGCATCCAG 540  
 GTACGCGCAG AAGATGACGA TTGTTTATCA ACGTCATCTA GAAGAATTG CCCCACACAG 600  
 GAACACATAG CAGATAAATC TCTGCACTGG AGGGGCTCT CCTCCCTCG TCTGCACT 660  
 GGAGCTCTAA TCCCACGCGC TGGATGAGT GCAGAAATAG CCGCGCAGG TATTTGTAAA 720  
 GTAGAGCCCA GCAGCTACAC ATGTACAAT TGCAAAACAGC CATTCAACAG TGCAATGGTT 780  
 CTCTTGCAAC CGGCACAGAA CACTCATGGA TTAAGAATCT ACTTAGAAGG GGAACACGGA 840  
 AGTCCCTCTG CCGCGCGGGT TGGTATCCCT TCAGGACTAG GTGCAGAAATG TCCTTCCCAG 900  
 CCACTCTTCC AGTGAATTCA TATTGCAGAC AATAACCCCT TTAACCTGCT AAGAATAACA 960  
 GGATCAGTAT CGAGAGAGGC TTCCGCGCTG GCAGAAAGGG GCTTTCCACC CACTCCCCC 1020  
 CTGTTTAGTC CACCAACGAG ACATCACTTG GACCCCAACC GCATAGAGCG CCTGGGGGCG 1080  
 GAAGAAATGG CCTTGGCCAC CCATCACCGG AGTGCCCTTG ACAGGGTGTCT GGGGTGAAT 1140  
 CCAATGGCTA TGGAGCTTCC CGCCATGGAT TTCTCTAGGA GACTTAGAGA GCTGGCAGGG 1200  
 AACAGCTCTA GCGCACAGCT GTCCCAAGGC CGGCCAGGCT CTATGCAAGG GTTACTGCAA 1260  
 CCATTCCAGC CAGGTAGCAA GCGCGCTTC CTGGCGACGC CCGCCCTCCC TCCTCTGCAA 1320  
 TCCGCCCCTC CTCCCTCCCA GCGCCCGGTC AAGTCCAAGT CATGCGAGTT CTGCGGCAAG 1380  
 ACGTTCAAAT TTCAGAGCAA CCGCGCGCA CACCGCGGCA GCCACACGGG CGAGAAGCCC 1440  
 TACAAGTGCA ACCTGTGCGA CCAAGCGTGC ACCCAGGCCA GCAAGCTGAA GCGCCACATG 1500  
 AAGACGCACA TGCACAAAAT GTCCCCCATG ACGGTCAAGT CGAGCGAGCG TCTCTCCACC 1560  
 GCCAGCTCCC CGGAACCGGG CACCAGCGAC TTGGTGGGCA GCGCCAGCAG CGCGCTCAAG 1620  
 TCCGTGTGTG CCAAGTTCAA GAGCGAGAAC GACCCCAACC TGATCCCGGA GAACGGGGAC 1680  
 GAGGAGGAAG AGGAGGACGA CGAGGAAGAG GAAGAAGAGG AGGAAGAGGA GAGGAGGAG 1740  
 CTGACGGAGA GCGAGAGGGT GGACTAGCGC TTCGGGTGTA GCCTGGAGGC GCGCGGCCAC 1800  
 CACGAGAACCA GCTCGCGGGG CGCGGTGCTG GCGGTGGGCG ACGAGAGCGG CGCCCTGCCC 1860  
 GACGTATGTC AGGGCATGGT GCTCAGCTCC ATGCAGCACT TCAGCGAGGC CTTCCACAG 1920  
 GTCTGTGGCG AGAAGCATAA GCGCGGCCAC CTGGCGAGG CCGAGGGCCA CAGGGACACT 1980  
 TGCGACGAAG ACTCGGTGGC CGCGAGTGC GACCGCATAG ACGATGGCAC TGTTAATGTC 2040  
 CGCGGTGCT CCGCGGGCGA GTGCGCTCG GGGGGCTGT CCAAAAAGCT GCTGCTGGGC 2100  
 AGCCCAAGCT CGCTGAGCCC CTCTCTAAG CGCATCAAGC TCGAGAAGGA GTTCGACCTG 2160  
 CCGCGCGCCA CGATGCCCAA CACGAGAAC GTGTACTCGC AGTGGCTGCG CGGCTACGCG 2220  
 GCCTCAGGC AGCTCAAAGA TCCCTTCTT AGCTTCGGAG ACTCCAGACA ATCGCTTTT 2280  
 GCCTCTCTGT CGAGCACTC CTGCGAGAAC GGGAGCTTGC GCTTCTCCAC ACCCGCCGG 2340  
 GAGCTGGAGC GAGGATCTC GGGGCGCAGC GGCACGGGAA GTGGAGGGAG CACGCCCCAT 2400  
 ATTAGTGGTC GCGGCACGGG CAGGCCAGC TCAAAAGAGG GCAGACGCGG CGACACTTGT 2460  
 GAGTACTGTG GGAAGTCTT CAAGAATCTG AGCAATCTCA CTGTCCACAG GAGAAGCCAC 2520  
 ACGGCGGAAA GGCCTTATAA ATGCGAGCTG TGCAACTATG CCGTGTGCCA GAGTAGCAAG 2580  
 CTCACCAAGC ACATGAAAC GCATGGCCAG GTGGGGAAGG ACGTTTACAA ATGTGAAAT 2640  
 TGTAAAGTGC TTTTAGCGT GTACAGTACC CTGGAGAAAC ACATGAAAAA ATGGCACAGT 2700  
 GATCGAGTGT TGAATAATG TATAAAACT GAATAGAGGT ATATTAAAC CCTTCCCTCA 2760  
 CTCCCACTG ACACCCCTT TTTCAACACT CCCTTTCCCC ATCGCCCTCC AGCCCCACTC 2820  
 CTCGTAGGAT TTTTCTAG TCCCATGTGA TTTAAACAAA CAAACAAACA AACAGAAAT 2880  
 ACGAAGCTAA GAAATATAGA GTGCTTGTCA CCAGCACACC TGTTTTTTTT CTTTTCTTT 2940  
 TTTTTTTTTC TTTTTCCTTT TTTTTTTTTT TCCTTTATGT TCTCACCGTT TGAATGCAAT 3000

	ATCTGTATGG	GGCAATACTA	TTGCAITTTA	CGCAAACTTT	GAGCCCTTCT	CTTGTGCAAT	3060
	AATTTACATG	TTGTGTATGT	TTTTTTTAA	ACTTAGACAG	CATGTATGGT	ATGTTATGGC	3120
	TATTTTAAAT	TGTCCTTAAT	TCGTTGCTGA	GCAAAACATG	TGCTGTTTCC	AGTTCCGTTT	3180
5	TGAGAGAAAA	AGAGAGAGAG	AGAGAAAAAG	ACCATGCTGC	ATACATTCTG	TAATACATAT	3240
	CATGTACAGT	TTTATTTTAT	AACGTGAGGA	GGAAAAACAG	TCTTTGGATT	AACCCCTCTAT	3300
	AGACAGAAAT	GATAGCACTG	AAAAAAATC	TCTATGAGCT	AAATGTCTGT	CTCTAAAGGG	3360
	TTAAATGTAT	CAATTTGGAA	GGAAGAAAAA	AGGCCCTTGA	TTGACAAATT	AACAGAAAAA	3420
	CAGAACAAAT	TTATTTCTAT	ATTTGGTTTT	AAAATATGAG	TGCCCTGGAT	CTATTAAAAC	3480
	CACATCGATG	GTTCCTTCTA	CTTGTATATA	ACTTGTAGCT	TAATTCAGCA	TTGGGTGAGG	3540
10	TAATAAACCT	TAGGAACCTAG	CATATAATTC	TATATTGTAT	TTCTCACAAAC	AATGGCTACC	3600
	TAAAAAGATG	ACCCATTATG	TCCTAGTTAA	TCATCATTIT	TCCTTTAGTT	TAATTTTATA	3660
	AACAAACTG	ATTATACCAG	TATAAAGCT	ACTTTGCTCC	TGGTGAGAGC	TTAAAAAGAA	3720
	TGGGCTGTTT	TGCCCAAGT	TTTATTTTTT	TTAAACAATG	ATTAATATGA	ATGTGTAAATG	3780
	TGCAAAAGCC	CTGGAACCGA	ATTAATACA	CTAGTAAGGA	GTTCAITTTA	TGAAGATATT	3840
15	TGCTTTAATA	ATGCTTTTTT	AAAAAATCTG	GCACCAAAAG	AAATAGATCC	AGATCTACTT	3900
	GGTTGTCAAG	TGGCAATCA	AATGATAAAC	TTTAAAGACCT	TGTATACCAT	ATTGAAAGGA	3960
	AGAGGCTGAC	AATAAGGTTT	GACAGAGGGG	AACAGAAGAA	AATAATATGA	TTTATTAGCA	4020
	CAACGCTGTA	CTATTGGCCA	TTTAAACTA	GAACAGGTAT	ATAAGCTAAT	ATTGATACAA	4080
	TGATGATTAA	CTATGAATTC	TTAAGACTTG	CATTATAAATG	TGACATTCIT	AAAAAAGAA	4140
20	GAGAAAGAA	TTTAAAGATA	GCAGTATATA	TGCTGTGCT	CCCTAAAAGT	TGTACTTCAT	4200
	TTCTTTTCCA	TACACTGTGT	GCTATTTGTG	TTAACATGGA	AGAGGATTC	TTGTTTTTAT	4260
	TTTATTTTTT	TTAATTTTTT	CTTTTTTAIT	AAGCTAGCAT	CTGCCCCAGT	TGGTGTCTAA	4320
	ATAGCACTTG	ACTTGCCTG	TGATATCTGT	ATCTTTTCTC	TAATCAGAGA	TACAGAGGTT	4380
	GAGTATAAAA	TAAACCTGCT	CAGATAGGAC	AATTAAGTGC	ACTGTACAAT	TTTCCAGTT	4440
25	TACAGGTCTA	TACTTAAGGG	AAAAGTTGCA	AGAATGCTGA	AAAAAAATG	AACACAATCT	4500
	CATTGAGGAG	CATTTTTTAA	AACTAAAAA	AAAAAAACT	TTGCCAGCCA	TTTACTTGAC	4560
	TATTGAGCTT	ACTTACTTGG	ACGCAACATT	GCAAGCGCTG	TGAATGGAAA	CAGAATACAC	4620
	TTAACAATGA	AATGAATGAT	TGCTTTCGCT	TCTACAGTGC	AAGGATTTTT	TTGTACAAAA	4680
	CTTTTTTAAA	TATAATGTTT	AAGAAAAATT	TTTTTTAAAA	AACACTTCAT	TATGTTTAGG	4740
30	GGGGAAGTGC	ATTTTAGGGT	TCCATTGTCT	TGGTGGTGT	ACAAGACTTG	TTATCCATT	4800
	AAAAATGGTA	GTGGAAATTC	TATGCCCTTG	ATACACACCG	CTCTTCAGGT	TGTAAAAAAA	4860
	AAAAACATAC	ATTGGGGAAA	GGTTTAAGAT	TATATAGTAC	TTAAATATAG	GAAAAATGCAC	4920
	ACTCATGTTG	ATTCCTATGC	TAAAAATACAT	TTATGGTCTT	TTTTCTGTAT	TTCTAGAAATG	4980
35	GTATTTGAAT	TAAATGTTCA	TCTAGTGTTA	GGCACTATAG	TATTTATATT	GAAGCTTGTA	5040
	TTTTTAACTG	TGCTGTGCT	TCTTAAAGG	TATCAATGTA	CCTTTTTTGG	TAGTGGAAAA	5100
	AAAAAAGACA	GGCTGCCACA	GTATTTTTTT	TTAATTGGC	AGGATAATAT	AGTGCAAATT	5160
	ATTTGTATGC	TTCAAAAAA	AAAAAAGAG	AGAAACAAA	AAGTGTGACA	TTACAGATGA	5220
	GAAGCCATAT	AATGGCGGTT	TGGGGGAGCC	TGCTAGAATG	TCACATGGAT	GGCTGTCTATA	5280
40	GGGGTTGTAC	ATATCCTTTT	TTGTTCCTTT	TTCTGCTGTC	CATACTGTAT	GCAGTACTGC	5340
	AAGCTAATAA	CGTGGTTTGT	TTATGTAGTG	TGCTTTTTGT	CCCTTTCCCT	CTATCACCTT	5400
	ACATTCAGAC	ATCTTACCTT	CATATGCAGT	AAAAGAAAGA	AAGAAAAAAA	AAGGAAAAAA	5460
	AAAAAANAAC	CAATGTTTTG	CAGTTTTTTT	CATTGCCAAA	AACTAAATGG	TGCTTTATAT	5520
	TTAGATTGGA	AAGAATTTCA	TATGCAAAAG	ATATTAAAGA	GAAAGCCCGC	TTTAGTCAAT	5580
45	ACTTTTTTGT	AAATGGCAAT	GCAGAAATAT	TTGTTATGG	CCTTTTCTAT	TCCTGTAAATG	5640
	AAAGCTGTTT	GTGTAACCTT	GAATTTTAT	CTTTTACTAT	GGGAGTCACT	ATTTATTATT	5700
	GCTTATGTGC	CTGTGTCAAA	ACAGAGGCAC	TTAATTGTAT	CTTTTATTTT	TCTTTGTTTT	5760
	TATTTTTTTT	TTTATTAGAA	TGACCAAAAG	TCATTACAAC	CTGGCTTTTT	ATTGTATTGT	5820
	TTCTGGTCT	TTGTAAGTT	CTATTGGAAG	AACCACTGTC	TGTGTTTTTT	TGGCAGTTGT	5880
50	CTGCATTAA	CTGTTTATAC	ACCCATTTTG	TCCTTTTATT	GAAAAAATAA	AAAAAATTA	5940

Seq ID NO: 85 Protein sequence:  
Protein Accession #: NP\_075044.1

55	1	11	21	31	41	51	
	MSRRKQKQKQ	HLSKREFSPE	PLEAILTDDE	PDHGPLGAPE	GDHLLTTCGQ	QCMNFPPLGDI	60
	LPIFIEHKRKQ	CNGSLCLEKA	VDKPPSPSP	EMKASNPVE	VGIQVTPEDD	DCLSTSSRRI	120
60	CPKQSHIADK	LLHWRGLSSP	RSAGHALIPT	PGMSAEYAPQ	GICKDEPSSY	TCTTCKQFFT	180
	SAWFLQLHAQ	NTHGLRIYLE	SEHSGPLTFR	VGIPSGLGAE	CPSQPPLBGI	HLADNNPFNL	240
	LRIPGSVSRQ	ASGLAEGFRP	PTPPLFSPPF	RHLDLPHRIE	RLGAEEMALA	THHPSAFDRV	300
	LRLNPMAMEP	PAMDFSRRLR	ELAGNTSSSP	LSPGRPSPMQ	RLQLPFQPGS	KPPFLATPPL	360
	PPLQSAFPSP	QPPVKSKECE	FCGKTFKFQS	NLVVHRSHT	GEKPYKCNLC	DHACTOASKL	420
65	KRHKMTHMK	SSPMVTKSD	GLSTASSPEP	GTSDLVGSAS	SALKSVVAKF	KSENDPNLIP	480
	ENGDEEREED	DEEEEREED	EEELTESER	VDYGFGLSLE	AARHHENSR	GAUVGVGDES	540
	RALPDVMQGM	VLSSMQHFSE	APHQVLGEKE	KRGHLAEAEQ	HRDTCDEDSV	AGESDRIDDG	600
	TVNGRGCSGP	ESASGGLSKK	LLLGSPSSLS	PFSKRIKLEK	EPDLPPATMP	NTENVYSQWL	660
	AGYAAASRLK	DPFLSFGDSR	QSPFASSSEH	SSENGSLRFS	TPPGELDGGI	SGRSGTGSGG	720
70	STPHISGFGT	GRPSSKEGRR	SDTCEYCGKV	FKNCSNLTWH	RRSHTGERPY	KCELCNYACA	780
	QSSKLTRHMK	THGQVKGQVY	KCBICKMPFS	VYSTLEKEMK	KWESDRVLNN	DIKTE	

Seq ID NO: 86 DNA sequence  
Nucleic Acid Accession #: XM\_035292.2  
Coding sequence: 53-1576

75	1	11	21	31	41	51	
	GCTCGCTGGG	CCGCGGCTCC	CGGGTGTCCC	AGGCCCGGCC	GGTGCGCAGA	GCATGGCGGG	60
80	TGCGGGCCCG	AAGCGGCGCG	CGCTAGCGGC	GCCGCGCGCC	GAGGAGAAGG	AAGAGGCGCG	120
	GGAGAAGATG	CTGCGCCCCA	AGAGCGCGGA	CGGCTCGCGC	CCGCGAGGCG	AGGGCGAGGG	180
	CGTGACCCCT	CAGCGGAACA	TCAAGCTGCT	CAACGCGGTG	GCCATCATCG	TGGGGACCAT	240
	TATCGGCTCG	GGCATCTTCT	TGACGCCAC	GGGCGTGCTC	AAGGAGGCG	GCTCGCGGGG	300
85	GCTCGGCGCT	GTGGTGTGGG	CCGCGTGCGG	CGTCTTCTCC	ATCGTGGGCG	CGCTCTGCTA	360
	CGCGGAGCTC	GGCACCAACA	TCTCACAATC	GGGCGGCGAC	TACGCTACCA	TGCTGGAGGT	420
	CTACGGCTCG	CTGCGCGCCT	TCCTCAAGCT	CTGGATCGAG	CTGCTCATCA	TCCGGCCTTC	480
	ATCGCAGTAC	ATCGTGGCCC	TGGTCTTCGC	CACCTACCTG	CTCAAGCCSC	TCTTCCCCAC	540

CTGCCCGGTG CCGGAGGAGG CAGCCAAGCT CGTGGCCTGC CTCTGCTGTC TGCTGCTCAC 600  
 GGCGGTGAAC TGCTACAGCG TGAAGGCCGC CACCCGGGTC CAGGATGCCT TTGCCGCGGC 660  
 CAAGCTCCTG GCCCTGGCCC TGATCATCCT GCTGGGGTTC GTCCAGATCG GGAAGGGTGA 720  
 TGTGTCCAAT CGCATCCCA ACTTCTCATT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780  
 TGTGTGGCA TTATACAGCG GCCTCTTTCG CTATGGAGGA TGAATTACT TGAATTTGCT 840  
 CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTCG GCCATCATCA TCTCCCTGCC 900  
 CATCGTGACG CTGGTGTACG TGCTGACCAA CCTGGCCTAC TTCACCAACC TGTCCACCGA 960  
 GCAGATGCTG TCGTCCGAGG CCGTGGCCGT GGACTTCGGG AACTATCACC TGGGCGTCAT 1020  
 GTCCCTGGAT ATCCCGCTCT TCGTGGCCCT GTCCCTGCTTC GGCTCCGTCA ATGGGTCCCT 1080  
 GTTACATCC TCCAGGCTCT TCTTCGTGGG GTCCCGGGAA GGCCACCTGC CCTCCATCCT 1140  
 CTCCATGATC CACCCACAGC TCCTCACCCC CGTGGCGTCC CTGCTGTTC CBTGTGTGAT 1200  
 GACGCTGCTC TAGCCCTTCT CCAAGGACAT CTCTCTCGTC ATCAACTTCT TCAGCTTCTT 1260  
 CAACTGGCTC TGGTGGCCCC TGGCCATCAT CGGCATGATC TGGCTGGGCC ACAGAAAGCC 1320  
 TGAGCTTGAG CGGCCATCA AGGTGAACCT GGCCCTGCCT GTGTTCCTCA TCCTGGCCTG 1380  
 CCTCTTCTG ATCGCCCTCT CCTTCTGGAA GACACCCGTG GAGTGTGGCA TCGGCTTTCAC 1440  
 CATCATCTC AGCGGGCTGC CGGTCTACTT CTTCGGGGTC TGGTGGAAAA ACAAGCCCAA 1500  
 GTGGCTCTC CAGGGCATCT TCTCCAGGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGCT 1560  
 CCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 87 Protein sequence:  
Protein Accession #: XP\_035292.2

1 11 21 31 41 51  
 | | | | |  
 MAGAGPKRRA LAAPAAEKE EAREKMLAAK SADGSAPAGE GEGVTLQRNI TLINGVAIV 60  
 GTIIGSGIFV TPTGVLEKAG SPGLALVVWA ACGVFSIVGA LCYAELETTI SKSGGDYAYM 120  
 LSVVSGSLPAF LKLEWIELII RPSSQYIVAL VFATYLLKPL PFTCPVPEEA AKLVACLCLVL 180  
 LLTAVNCYSV KAATRVQDAF AAKLLALAL IILLGFVQIG KQDVSNLDPN PSFEGTKLDV 240  
 GNVILALYSG LFATYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVVV LTNLAYPTTL 300  
 STEQMLSEEA VAVDFGNYHL GVMWIIIPVF VGLSCFPGSVN GSLPTSSRLF PVSGREGHLP 360  
 SILSMIHQPL LTPVPSLVFT CVMTLIYAFS KDIPSVINFP SFFNWLCLAL AIIGMIWLRLH 420  
 RKPELERPIK VMLALPVFFI LACLELIAVS FWKTFVECGI GFTIILSGLP VYFPGVWKNY 480  
 KPKWLLQGIF STTVLCQKLM QVVPQET

Seq ID NO: 88 DNA sequence  
Nucleic Acid Accession #: NM\_005268.1  
Coding sequence: 168-989

1 11 21 31 41 51  
 | | | | |  
 TAAAAAGCAA AAGAATTGCG GGCGCGCTCG ACACGGGCTT CCCCAGAAAC CTTCCTCCGCT 60  
 TCTGATATG AAATCTAAGC TGCTTGTGTA GTCTATTGCG CGGCTGCTGG GAGCCAGGAG 120  
 AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACCGGTGG GTCCACCATG AACTGGAGTA 180  
 TCTTTGAGGG ACTCTGAGT GGGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC 240  
 TGTCTCTGGT TTTCATCTTC CGCGTGTGCG TGACTCTGGT GACGGCCGAG CGTGTGTGGA 300  
 GTGATGACCA CAAGGACTTC GACTGCAATA CTGCGCCAGC CGGCTGCTCC AACCTCTGCT 360  
 TGTATGAGTT CTTCCTGTGT TCCCATGTGC GCCTCTGGGC CCTGCAGCTT ATCCTGGTGA 420  
 CATGCCCTC ACTGCTGTGT GTCATGCACG TGGCTTACCG GGAGGTTCAG GAGAAGAGGC 480  
 ACCGAGAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG 540  
 GTGGGCTCTG GTGGACATAT GTCTGCAGCC TAGTGTTCAG GCGAGAGGTG GACATCGCCT 600  
 TTCTCTATGT GTTCCACTCA TTCTACCCCA AATATATCCT CCCTCCTGTG GTCAAGTGCC 660  
 ACGCAGATCC ATGTCCCAAT ATAGTGGACT GCTTCACTTC CAAGCCCTCA GAGAAGAAAC 720  
 TTTTCACTCT CTTCATGGTG GCCACAGCTG CCATCTGCAT CCTGCTCAAC CTGCTGGAGC 780  
 TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCCTGGC AGCAAGGAAA GCTCAAGCCA 840  
 TGTGCACAGG TCATCACCCC CACGGTACCA CCTCTTCCTG CAACAAGACG GACCTCTCTT 900  
 GGGGTGACCT CATCTTCTG GGCTCAGACA GTCACTCTCC TCTCTTACCA GACCGCCCCC 960  
 GAGCAATATG GAAGAAAAAC ATCTTGTGAG GGGCTGCTGT GACTGGTCTG GCAGGTGGG 1020  
 CCTGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC 1080  
 CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGACGCTC TGGGAGCCAG TTCCTAGTCC 1140  
 TCAACTCCAG CCACCTGCCC CAGCTCGACG GCATCGGGCC AGTTCCCCCT CTGCTCTGCA 1200  
 GCTCGGTTTC CTTTCTAGA ATGGAAATAG TGAGGGCCAA TGC

Seq ID NO: 89 Protein sequence:  
Protein Accession #: NP\_005259.1

1 11 21 31 41 51  
 | | | | |  
 MNWSIFEGLL SGVNKYSTAF GRIWLSLVFI FRVLVYLVT A ERVWSDHDKD FDCNTRQPGC 60  
 SNVCFDEFFP VSHVRLWALQ LILVTCPSLL VVMHVAYREV QEKREAREAHG ENSGRLYLNP 120  
 GKRRGGIWNV YVCSLVFKAS VDIAPLYVFH SFYPKYILPP VVKCHADPCP NIVDCFLSKP 180  
 SEKNIFTLFM VATAAICILL NLVELIYLV S KRCHCELAAR KAQAMCTGHE PHGTTSSCKQ 240  
 DDLLSGDLIF LGSDSHPPLL PDRPRDHVKK TIL

Seq ID NO: 90 DNA sequence  
Nucleic Acid Accession #: NM\_002391.1  
Coding sequence: 26-457

1 11 21 31 41 51  
 | | | | |  
 CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCACCCTCCT 60  
 CGCCCTGCTG GCGCTCACCT CGCGGTGCG CAAAAGAGAA GATAAGGTGA AGAAGGGGCG 120  
 CCGGGGAGC GAGTGGCTG AGTGGGCTG GGGGCGCTGC ACCCCAGCA GCAAGGATTG 180  
 CGGCGTGGT TTCCGCGAGG GCACCTGCGG GGGCCAGACC CAGCGCATCC GGTGCAGGGT 240  
 GCCCTGCAAC TGGAGAGAGG AGTTTGGAGC CGACTGCAAG TACAAGTTTG AGAACTGGGG 300  
 TCGTGTGAT GGGGCGACAG GCACCAAGT CGCCCAAGGC ACCCTGAAGA AGGGCGGCTA 360

CAATGCTCAG TGGCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCCA AGACCAAGGC 420  
 AAAGGCCCAA GCGAAGAAAG GGAAGGGAAA GGACTAGACG CCAAGCCTGG ATGCCAAGGA 480  
 GCGCCCTGGT TCACATGGGG CCTGGCCACG CCCTCCCTCT CCCAGGCCCG AGATGTGACC 540  
 CACCAAGTGC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCTGCGCT TGTCCCTCTC 600  
 ACTCCCCAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660  
 TGAGCCTCCC CCAAGCAAT GTGAGTCCCA GAGCCGCGTT TTGTTCTTCC CCACAATTCC 720  
 ATTACTAAGA AACACATCAA ATAAACTGAC TTTTCCCCC CAATAAAGC TCTTCTTTT 780  
 TAATAT

Seq ID NO: 91 Protein sequence:  
 Protein Accession #: NP\_002382.1

1 11 21 31 41 51  
 MQHRGFLLLT LLALLALTS VAKKKDKVKK GPGSECAEW ANGPTPSSK DGVGFREGT 60  
 CGAQQRIRC RVPCNWKKEP GADCKYFEN WGACDGGTGT KVRQGLTKKA RYNAQCQETI 120  
 RVTKPCFPKT KAKAKAKKKG GKD

Seq ID NO: 92 DNA sequence  
 Nucleic Acid Accession #: NM\_005130.1  
 Coding sequence: 98-802

1 11 21 31 41 51  
 CTCTACCTGA CACAGCTGCA GCCTGCAATT CACTCCCACT GCCTGGGATT GCACTGGATC 60  
 CGTGTGCTCA GAACAAGTGT AACGCCACG TGCAGCCATG AAGATCTGTA GCCTCACCCCT 120  
 GCTCTCCTTC CTCTACTGG CTGCTCAGGT GCTCTCTGGT GAGGGGAAAA AAAAAGTGAA 180  
 GAATGGACTT CACAGCAAAG TGGTCTCAGA ACAAAAGGAC ACTCTGGGCA ACACCCAGAT 240  
 TAAGCAGAAA AGCAGGCCCG GGAACAAAGG CAAGTTTGTC ACCAAAGACC AAGCCAACTG 300  
 CAGATGGGCT GCTACTGAGC AGGAGGAGGG CATCTCTCTC AAGGTTGAGT GCACTCAATT 360  
 GGACCATGAA TTTCTCTGTG TCTTTGCTGG CAATCCAACC TCATGCCATA AGCTCAAGGA 420  
 TGAGAGAGTC TATTGGAAAC AAGTTGCCCG GAATCTGOGC TCACAGAAAG ACATCTGTAG 480  
 ATATTCCAAG ACAGCTGTGA AAACCAAGAT GTGCAGAAAG GATTTTCCAG AATCCAGTCT 540  
 TAAGCTAGTC AGCTCCACTC TATTGGGAA CACAAAGCCC AGGAAGGAGA AAACAGAGAT 600  
 GTCCCCCAGG GAGCACATCA AGGGCAAAGA GACCACCCCT TCTAGCCTAG CAGTGACCCA 660  
 GACCATGGCC ACCAAAGCTC CCGAGTGTGT GGAGGACCCA GATATGGCAA ACCAGAGGAA 720  
 GAGTGCCTGT GAGTCTCTGT GAGAGACTTG GAGCTCTCTC TGACATCTCT TCCTCAGCAT 780  
 AGTGAGAGAC AGCTCATGCT AATGAGGTCA AAAGAGAACG GGTTCCTTTA AGAGATGTCA 840  
 TGTGCTAAGT CCTCTGTAT ACTTTAAAGC TCTCTACAGT CCCCCAAAAA TATGAACCTT 900  
 TGTGCTTACT GAGTGCAACG AATATTAA ACAAGTTTGT TATTTTGTGC TTTTGTGTTT 960  
 TGGAAATTTG CTTAATTTTC TTGGATGCGA TGTTCAGAGG CTGTTTCTGT CAGCATGTAT 1020  
 TTCCATGGCC CACACAGCTA TGTGTTTGTG CAGCGAAGAG TCTTTGAGCT GAATGAGCCA 1080  
 GAGTGATAAT TTAGTGCAA CGAACTTTCT GCTGAATTAA TGGTAATAAA ACTCTGGGTG 1140  
 TTTTCAAAA AAAAAAAAAA AAA

Seq ID NO: 93 Protein sequence:  
 Protein Accession #: NP\_005121.1

1 11 21 31 41 51  
 MKICSLTLLS FLLAAQVLL VEGKKKVKNG LHSKVVSEQK DTLGNTQIKQ KSRPGNKQKP 60  
 VTQDQANCRW AATEQEGBIS LKVECTQLDH EFSCVFAGNP TSCLKLKDPR VYWKQVARNL 120  
 RSQKIDICRY S KTAVKTRVCR KDFPSSSLKL VSSTLFGNTK PRKERTMSP REHIGKETT 180  
 PSSLAVTQTM ATKAPECEVD PDMANQRKTA LEPCGWTSS LCTFFLSIVQ DTSC

Seq ID NO: 94 DNA sequence  
 Nucleic Acid Accession #: NM\_012101  
 Coding sequence: 125-1891

1 11 21 31 41 51  
 CTCTCACAG GTGTGTCTCT AGTCTCTGTG GTTGCCTGCC CCACTCCCTG CCGAGACGCC 60  
 TGCCAGAAAG GTCACCTATC CTGAACCCCA GCAAGCCTGA AACAGCTCAG CCAAGCACCC 120  
 TGGATGGAA GCTGCAGATG CCTCCAGGAG CAACGGGTCTG AGCCCAAGAG CCAGGGATGC 180  
 CCGGAGCCCG TCGGGCCCA GTGGCAGCCT GGAGAATGGC ACCAAGGCTG ACGGCAAGGA 240  
 TGCCAGAAC ACCAACGGGC ACGGCGGGGA GGCAGCTGAG GGCAGAGGCC TGGGACGGC 300  
 CTGAAGCCA GGGGAAGGTA GGAGCGCCCT GTTCGCGGGC AATGAGTGGC GCGACCCAT 360  
 CATCCAGTTT GTCAGTCCG GGGACGACAA GAACTCCAAC TACTTCAGCA TGGACTCTAT 420  
 GGAAGGCAAG AGGTGCGCGT ACAGGAGGCT CCAGCTGGGG GCTGCCAAGA AGCCACCCGT 480  
 TACCTTTGCC GAAAGGGGCG ACGTGCGCAA GTCCATTTTC TCGAGTCCC GGAAGCCAC 540  
 GGTGTCCATC ATGAGACCCG GGGAGACCCG GCGGAACAGC TACCCCGGGG CCGACACGG 600  
 CCTTTTTC CAAGTCCAAGT CCGGCTCCGA GGAGGTGCTG TGCGACTCCT GCATCGGCAA 660  
 CAGCAGAAAG GCGGTCAAGT CCGGCTCCGA GTGCCAGGCC TCCTTCCTGG AGCTGCATCT 720  
 CAGCCCCAC CTGAGGCGCG CCGGCTCCGA AGACCCAG CTGCTGAGC CCATCCGGGA 780  
 CTTTGAAGCC GCGAAGTGT CCGTGCATGG CAAGACGATG GAGCTCTTCT GCCAGACCGA 840  
 CCAGACCTGC ATCTGTCTACC TTTGCATGTT CCAGGAGCAC AAGAATCATA GCACCGTGAC 900  
 AGTGGAGGAG GCGAAGGCG AGAAGGAGAC GGAGCTGTCA CTGCAAAAGG AGCAGCTGCA 960  
 GCTCAGATG ATTGAGATTG AGGATGAAGC TGAGAAGTGG CAGAAGGAGA AGGACCGCAT 1020  
 CAAGAGCTTC ACCCAATG AGAAGGCCAT CCGGAGCAG AACTTCCGGG ACCTGCTGCG 1080  
 GGACCTGGAG AAGCAAAAG AGAAGTGAG GGCTGCGCTG GAGCAGCGGG AGCAGGATGC 1140  
 TGTGGACCA GTGAAGTGA TCATGGATGC TCTGGATGAG AGAGCCAAG TGTGTCATGA 1200  
 GGACAGCAG ACCCGGAGC AGCTGCATAG CATCAGGAC TCTGTGTTGT TCTGCGAGA 1260  
 ATTTGGTGA TTGATGAGA ATTACTCTCT CCCCCACCC CTGCCCACCT ATCATGTCT 1320  
 GCTGGAGGG GAGGCGCTCG GACAGTCACT AGGCAACTTC AAGGACGACC TGCTCAATGT 1380  
 ATGCATGCGC CAGCTTGAAG AGATGTGCAA GCGGACCTG AGCCGTAAC TCATTGAGAG 1440  
 GAACACATG GAGAACGGTG GTGACCATCG CTATGTGAAC AACTACACGA ACAGCTTCGG 1500

GGGTGAATGG AGTGACACCG ACACCATGAA GAGATACTCC ATGTACCTGA CACCCAAAGG 1560  
 TGGGGTCCGG ACATCATACC AGCCCTCGTC TCCTGGCCGC TTCAACCAAGG AGACCACCCA 1620  
 GAAGAATTTT AACAACTCTT ATGGCACCAA AGGTAACATC ACCTCCCGGG TCTGGGAGTA 1680  
 CTCCTCCAGC ATTCAGAACT CTGACAAATGA CCTGCCCGTC GTCCAAGGCA GCTCCTCTTT 1740  
 CTCCTGAAA GGCTATCCCT CCCTCATGCG GAGCCAAAGC CCCAAGGCC AGCCCCAGAC 1800  
 TTGGAAATCT GGCAGACAGA CTATGCTGTC TCACTACCGG CCATTCTACG TCAACAAAGG 1860  
 CAACGGGATT GGGTCCACAG AAGCCCCATG AGCTCCTGGC GGAAGGAAGC AGGCGCCACA 1920  
 CCCCTGCTCT TCCTCTGAC CCTGCTGCTC TTGCCCTTCTA AGCTACTGTG CTGTGCTGGG 1980  
 TGGGAGGGAG CTTGCTCTG CACCTGCCCT CTGCAGCCCT CTGCCAGCCT CTGGGGGCA 2040  
 GTTCCGGCCT CTCGACTTC CCCACTGGCC ACACCTCCAT CAGACTCCTT TCCTGCTTTG 2100  
 TGACCTCAGA TGTTCACCAT CATTCCTGTG CTCAGAGGCC AACCCATCAC AGGGGTGAGA 2160  
 TAGGTTGGGG CCTGCCCTAA CCGCCAGGCC TCCTCCTCTC GGGCTGGATG TGGGGGCTAG 2220  
 CAGTGAGTAC CGCATGGTA TCAGCTGCTC TCCTCCGCCC ACGCCCTGCT GTCTCCAGGC 2280  
 CTATAGACGT TTCTCTCAA GGCCTATACC CCCAATGTTG TCAGCAGATG CCTGGACAGC 2340  
 ACAGCCACCC ATCTCCCAT CACATGGCCC ACCTCCTGCT TCCCAGAGGA CTGGCCCTAC 2400  
 GTGCTCTCTC TGCTCTTACC TATCAATGCC CAGCATGGCA GAACCTGCAG TGGCCAAAGG 2460  
 CTGCAGATGG AAACCTCTCA GTGCTTGAC ATCACCCTAC CCAGGGGGTG GGTCTCCACC 2520  
 ACAGCCACTT TGAGTCTGTG GTCCCTGGAG GGTGCTTCT CTGACTGGC AGGATGACCT 2580  
 TAGCCAAGAT ATTCTCTGT TCCCTCTGCT GAGATAAAGA ATTCCTTAA CATGATATAA 2640  
 TCCACCATG CAAATAGCTA CTGGCCAGC TACCATTAC CATTTGCTA CAGAATTTCA 2700  
 TTCAGTCTAC ATCTTGGCAT TCTCTCTGGC GATGGAGTGT GGCTGGGCTG ACCGCAAAAG 2760  
 GTGCTTACA CACTGCCCCC ACCCTCAGCC GTTGCCCATC CAGAGGCTGC CTCTCCTTC 2820  
 TGATTACCCC CCATGTTGCA TATCAGGGTG CTCAGGATT GGAGAGGAGA CAAACACAGG 2880  
 AGCAGACAG TGGGGACATC TCCGCTCTCA ACAGCCCGAG GCCTATGGG GCTCTGGAAG 2940  
 GATGGGCCAG CTTCGACGGG TTGGGGAGGG AGACATCCAG CTGGGCTTT CCCCTTTGGA 3000  
 ATAAACCATT GGTCTGTC

Seq ID NO: 95 Protein sequence:  
 Protein Accession #: NP\_036233.1

1 11 21 31 41 51  
 MEAADASRSN GSSPEARDAR SPSGPSGSLE NGTKADKDA KTTNGHGEA AEGKSLGSAL 60  
 KPEGERSALF AGNENRPII QFVSGDDKN SNYFSMDME GKRSFYAGLQ LGAARKPPVT 120  
 FAKKDVRSK IFSESRRPTV SIMEFGETRR NSYPRADTGL FSRKSGSEE VLCDSCIGNK 180  
 QKAVKSLVC QASFCBHLK PHLEGAARFD HQLLEPIRDF EARKCPVHGK TMELFCQTDQ 240  
 TCICYLCMPQ EKHHSFTVTV EKAKEKETE LSLQKBLQL KIIIEIDEAE KWQKEKDRIK 300  
 SFTTNEKAIL BQNFRLVRD LEKQKEEVRA ALEQRBDQAV DQVKVIMDAL DERAKVLHED 360  
 KQTRQLHSI SDSVLFLQEF GALMSNYSLP PPLPTYHVL EGEGLQSLG NFKIDLLNVC 420  
 MRHVEKCKA DLSRNFIERN HMENGDDERY VNNYNSFGG EWSAPDTMKR YSMYLPKGG 480  
 VRTSYQPSPP GRFTKETQK NFNLYGTGK NYTSRVWEYS SSIQNSDNDL PVVQGSSEPS 540  
 LRGPYSLMRS QSPKAQPTW KSGKQTMLSH YRPPYVKNGN GIGSNBAP

Seq ID NO: 96 DNA sequence  
 Nucleic Acid Accession #: NM\_080668.1  
 Coding sequence: 83-841

1 11 21 31 41 51  
 GGCACGAGGG CAGCGAGTGG CCTTCCCGGT TGGCGCGCGC CCGGGGCGGC GCGGCTGGAG 60  
 GAGCTCGAGA CGGAGCCTAG TTATGCTTGG GAGGCGAAAG CGTCCCGGAG GAGCCGCTCA 120  
 CGCTCCGGG CCAAGGCGCC CATCTCCTAC TAAGCCTCTG CGGAGGTCCC AGCGGAAATC 180  
 AGGCTCTGAA CTCGCGAGCA TCCTCCCTGA AATCTGGCCG AAGACACCCA GTGCGGCTGC 240  
 AGTCAGAAAG CCCATCGTCT TAAAGAGGAT CGTGGCCCAT GCTGTAGAGG TCCACGCTGT 300  
 CCAATCACCT CGCAGGAGCC CTAGGATTTC CTTTTCTTGT GAGAAAGAAA ACGAGCCCCC 360  
 TGGCAGGAGC CTTACTAAGG AGGACCTTTT CAAGACACAC AGCGTCCCTG CCACCCCCAC 420  
 CAGCACTCCT GTGCGGAACC CTGAGGCCGA GTCCAGCTCC AAGGAAGGAG AGCTGGACGC 480  
 CAGAGACTTG GAAATGTCTA AGAAAGTCAG CGGTCTCTAC AGCCGGCTGG AGACCCCTGG 540  
 CTCTGCTCTT ACCTCCACCC CAGGCCCGCG GTCTGCTTTT GGCTTCGAGG GGCTGCTGGG 600  
 GGCAGAGAGC TTGTCCGAG TCTCGCCAGT GGTGTGCTCC AAACCTCACG AGGTCCCCAG 660  
 GGTGTTGTGA AAGCCCTGGG CCCCAGACAT GACTCTCCCT GGAATCTCCC CACCACCCGA 720  
 GAAACAGAAA CGTAAGAGA AGAAAATGCC AGAGATCTTG AAAACGGAGC TGGATGAGTG 780  
 GGCTGGCGCC ATGAATGCCG AGTTTGAAGC TGCTGAGCAG TTGTGCTCTC TGGTTGAATG 840  
 AGATGCAATG GGGGGTGCAC CTGGCCAGAC TCTCCTCTCT GTCTGTACA TAGCCACCTC 900  
 CCTGTGGAGA GGACACTTAG GGTCCCTCTC CCGTCTCTTG TTAACCTGTG GTGTGCTGGT 960  
 GCTGCGCATG AGGACTGTCT GCCTTTGAGG GCTTGGGCG CAGCGGCAGC CATCTTGGTT 1020  
 TTAGGAAATG GGGCGGCTCG GCCCAGCCAC TCACTGGTGT CCGTCTCTCT GTCTGCTGTG 1080  
 CCTTCTATC TCCCCAAAGT ACCATAGCCA GTTTCAGAT GGGCCACAGA CTGGGGAGGA 1140  
 GAATCAGTGG CCCAGCCAGA AGTTAAAGGG CTGAGGGTTG AGGTGAGAGG CACCTCTGCT 1200  
 CTGTGTGGGA GGGGTGGCTG CTTGGAAATA GGCCAGGGG CTCTGCCAGC CTGGCCCTCT 1260  
 CCTCTCTGAG TTGCCCTCTG TTGGTGGCTT TCTTCTTGAA CCCACCTGTG TAAAGAGGTT 1320  
 TTCAGTTCCG TGGGTTTCCC CTTTGATTCT GTAAATAGTC CCAGAGAGAA TTCGTGGGCT 1380  
 GAGGGCAATT CTGCTTGGGA GGAAGAAGCT GGACATTCAG CCGTGTGGAGT CTGAGTTTGT 1440  
 AAGGATGTAG GGAGCCTTAG TTGGGTCTCA GACCATAAAT GTGTACTACA CAGAAGCTGT 1500  
 GTTTCTAGT TCTGCTCTCG TGTGAGATG TTTGTAATG GCCAGGTGTA TAGGGCGCTG 1560  
 GCTGCTTGGG GCAAGGGGTG CATTTGAGG TGTGCCACAC AGGTGCTGTG AGTTTCTGTG 1620  
 GCTCATGGCC TCTGGGCTGG TCCCTTGCAC AGGGCCCAAG CTGGAGTCTT ACCACTCTGC 1680  
 TGCAGGGGTG GAAGGTGGCC CCTCTTGTCA CCCATACCCA TTTCTTACAA AATAAGTTAC 1740  
 ACCGAGTCTA CTGAGCCCTA GAAGAGAAAG TTGAAGAGTC CCAGACCTAC TAGCATTTTG 1800  
 CAACTATGCT TGTAAAGTCC TCGGAAAGTT TCCTCGGTA CCAGACAGCG GCGGGGGCTG 1860  
 ATAGCAATTT TAGTTTTTGG CCTCCCTATC CTCTCACATG AGAACACTGC CTGGATGCAT 1920  
 CTCATGATCT CTGGAGATT TCCCATCTT TCTCTCTTCT CCATGCTGTG GATTCAATAG 1980  
 TTTGATTTG AAGGCTGCCC TGCCCCGAC TCTCTGCGG CACCCCTGGC CATTTGACTT 2040  
 TTTGATTTT AGAAGTTCTG GGAAGTAGAC GCTGAGGTGT GCAGAGGAGC TGGTGGATAA 2100  
 CAGAGAAATG CAGGGAAGAT GAGTGTCTGG TCAGGGTACT TGGATGAAGC GGTGCAGGCC 2160  
 AGGCGGGCCC TAATAAAACC CTCTGCCAGG TCTGGAGTC CCAGGCCATC TGCTCAAGCC 2220



TCTGTGGTTT GTCAGACCTG CAAGCAAGCC CCTGCTGGG GAAGCCTAGG TGTCTTGAG 2280  
 CTGAACCGCA CTGAAGAACT CTTGTCTCA CTGGCTGATG CAGCAGAACT CTTGGGAAAT 2340  
 GTCTTAGTCC TGCAGAAATCA GGAGTCACCA GATGATGCAG AGTTGAGATC ATCATTGCAA 2400  
 AGTTCTCTGT TCCTGAGGAA CTAATTTAA GGAATAAATG GGATTTTGT TTAGAGTTGG 2460  
 5 AAAAAAGGCC TGATTAAGA GTTCTGCCT GTTAAAAAA AAAAAA AAAAA

Seq ID NO: 97 Protein sequence:  
 Protein Accession #: NP\_542399.1

10 1 11 21 31 41 51  
 MSGRRTRSGG AAQSGPRAP SPTKPLRRSQ RKSGSELPFI LPBIWPKTPS AAARVKPIVL 60  
 KRIVAHAVEV PAVQSPRRSP RISFFLEKEN EPPGRELTKE DLFKTHSVPA TPTSTPVFNP 120  
 15 EAESSSEKEG LDARDLEMSK KVRYSYRLE TLGSASTSTP GRRSCFGFEG LLGAEDLSGV 180  
 SPVVCCKLKE VPRVCAKPWA PDMTLPGISP PPEKQKRKKK KMPILKTEL DEWAAAMNAE 240  
 FEAAEQFDLL VE

Seq ID NO: 98 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 58-12444

20 1 11 21 31 41 51  
 25 GGGGCATTTC GGGGTCCGGG CCGAGCGGGC GCACGCGCGG GAGCGGGACT CGGCGGCATG 60  
 GCGGGCTCCG GAGCCGGTGT GCGTGTCTCC CTGCTGCGGC TGCAGGAGAC CTTGTCCGCT 120  
 GCGGACCGCT GCGGTGCTGC CCTGGCCGGT CATCAACTGA TCCGCGCCT GGGGAGGAA 180  
 TCGTCTCTGA GCAGCAGCCC CCGGTGCTG GCATTACAGA CATCTTAGT TTTTCCAGA 240  
 GATTTCCGTT TGCCTGTATT TGTCCGGAAG TCATCAACA GTATTGAAT TCGTGAATG 300  
 30 AGAGAAGAA TCCCTAAGTT TTTATGATT TTCTAGAAA AAATGGGCA GAAGATGCA 360  
 CCTACTCTG TTGAAATTA GAACACTTGT ACCAGTGTAT ATACAAAAGA TAGAGCTGCT 420  
 AAATGAAAA TTCCAGCCCT GGACCTCTT ATTAAGTTAC TTCAGACTT TAGAAGTTCT 480  
 AGACTCATGG ATGAATTTAA AATTGGAGAA TTATTAGTA AATTCTATG AGAAGTTGCA 540  
 TTGAAAAAA AAATACCAGA TACAGTTTAA GAAAAAGTAT ATGAGCTCCT AGGATTATTG 600  
 35 GGTGAAGTTC ATCCTAGTGA GATGATAAAT AATGCAGAAA ACCTGTTCCG CGCTTTCTG 660  
 GGTGAACCTA AGACCCAGAT GACATCAGCA GTAAGAGAGC CCAAACTACC TGTCTCGCA 720  
 GGATGCTGA AGGGGTGTC CTCACCTCTG TGCAACTTCA CTAAGTCCAT GGAAGAAGAT 780  
 CCCCAGACTT CAAGGAGAGT TTTAAATTT GTACTAAAG CAATTCTGTC TCAGATTGAT 840  
 CTGAAGAGAT ATGCTGTGCC CTCAGCTGGC TTGCGCTAT TTGCGCTGCA TGCATCTCAG 900  
 40 TTAGCACCT GCCCTCTGGA CAACTACGTG TCTCTATTG AAGTCTTGT AAAGTGGTGT 960  
 GCGCCACAA ATGTAGAAT GAAAAAGCT GCACCTTCAG CCTGGAATC CTTTCTGAA 1020  
 CAGGTTCTA ATGTGGTGGC GAAAAATGCA GAAATGCATA AAAATAAAT GCAGTACTTT 1080  
 ATGAGCAGT TTTATGGAAT CATCAGAAAT GTGGATTGCA ACAACAAGGA GTTATCTATT 1140  
 GCTATCCGTG GATATGGACT TTTTGCAAG CCGTGCAAG TTATAAAGC AAAAGATGTT 1200  
 45 GACTTCATG ACGTTGAGCT CATTGAGCG TGCAAGCAGA TGTCTCTCAC CCAGACAGAC 1260  
 ACTGGTACG ACCGTGTTA TCAGATGCCA AGCTTCTCC AGTCTGTTC AAGCGTCTG 1320  
 CTGTAACCTG ACACAGTTTC TGAGGTGTAT ACTCCAGTTC TGGAGCACCT CGTGGTGATG 1380  
 CAGATAGACA GTTCCACCA GTACAGTCCA AAAATGCAGC TGGTGTGTT CAGAGCCATA 1440  
 GTGAAGGTGT TCCTAGCTTT GGCAGCAAAA GGGCCAGTTC TCAGGAATTG CATTAGTACT 1500  
 50 GTGGTGATC AGGGTTAAT CAGAATATGT TCTAAACAG TGGTCTCTCC AAAGGCGCT 1560  
 GAGTCTGAAT CTGAAGACCA CCGTGCTTCA GGGGAAGTCA GAATCGGCA ATGGAAGGTG 1620  
 CCCACATACA AGCATACGT GGATCTCTC AGACATCTCC TGAGCTCTGA CCAGATGATG 1680  
 GATTCTATT TAGCAGATGA AGCATTTTC TCTGTGAAT CCTCCAGTGA AAGTCTGAAT 1740  
 CATTACTTT ATGATGAAT TGTAATAATC GTTTTGAAGA TTGTTGAGAA ATTGGATCTT 1800  
 55 ACACCTGAAA TACAGACTGT TGGGAACAA GAGAATGGAG ATGAGGCGCC TGGTGTGTTG 1860  
 ATGATCCCAA CTTAGATAGC AGCGGCTAAC TTGCATCCAG CTAAACCTAA AGATTTTTCG 1920  
 GCTTCTATTA ACCTGGTGGG ATTTGCGAGA GAGATTCTCC CTGAGAAACA AGCAGAAATT 1980  
 TTTGAACCAT GGTGTACTTC ATTTTCATAT GAATTAATTT TGCAATCTAC AAGGTTGCC 2040  
 60 CTCTCAGATG GTTCTACAA ATTGCTTCT ATTACAGTAA GAAATGCCAA GAAATAAATA 2100  
 TATTTCCAGG GAGTTAGTCC AAAGAGTCTG AACACTCTC CTGAAGACCC AGAAAGTAT 2160  
 TCTGTCTTG CTTTATTGT GAAATTTGGC AAAGAGGTGG CAGTTAAAT GAAGCAGTAC 2220  
 AAAGATGAAC TTTTGGCTC TTGTTGACC TTTCTCTGT CCTGCCACA CAACATCAT 2280  
 GAACCTGATG TTAGAGCCTA OGTTCTGCA CTGCAGATGG CTTTCAACT GGGCTGAGC 2340  
 TATACCCCTG TGGCAGAAGT AGGCTGAAT GCTCTAGAAG AATGGTCAAT TTATATTGAC 2400  
 65 AGACATGTAA TGCAGCCTTA TTACAAAGAC ATTCTCCCTC GCCTGGATGG ATACCTGAAG 2460  
 ACTTCAGCCT TGTCAAGTA GACCAAGAA AACTGGGAAG TGTGAGCTT TCTCGGCT 2520  
 GCGCAGAAAG GATTTAATAA AGTGGTGTTA AAGCATCTGA AGAAGACAAA GAACCTTCA 2580  
 TCAACGAAAG CAATATCCTT AGAAGAAATA AGAATTAGAG TAGTACAAAT GCTTGGATCT 2640  
 CTAGGAGGAG AAATAACAA AAATCTCTG ACAGTCAGT CCTCAGATGA GATGATGAAG 2700  
 70 AGCTATGTGG CTTGGGACAG AGAGAAGCGG CTGAGCTTGG CAGTGCCTT TAGAGAGATG 2760  
 AAACCTGTCA TTTTCTGGA TGTGTTCTG CCTCGAGTCA CAGAATTAGC GCTCAGACC 2820  
 AGTGACAGAC AAATAAAGT TGCAGCCTGT GAACTTTTAC ATAGCATGGT TATGTTTATG 2880  
 TTGGGCAAG CCACGAGAT GCCAGAAAGG GGACAGGGAG CCCCACCAT GTACCAGCTC 2940  
 TATAAGCGGA CGTTTCTGT GCTGCTTGA CTTGCGTGTG ATGTTGATCA GGTGACAAGG 3000  
 75 CAACTGTATG AGCCACTAGT TATGCACTG ATTCACTGGT TCACTAACAA CAAGAAATTT 3060  
 GAAAGTCAGG ATACTGTGTC CTTACTAGAA GCTATATTG ATGGAATTGT GGAACCTGTT 3120  
 GACAGTACTT TAAGAGATT TTTGGTCCG TGTATTCGAG AATTCCTTAA ATGGTCCATT 3180  
 AAGCAATAA CACACAGCA GCAGGAGAAG AGTCCAGTAA ACACCAATC GCTTTTCAAG 3240  
 CGACTTTATA ACCTTGCCTC TCACCCCAAT GCTTTCAAGA GGCTGGGAGC ATCACTTGCC 3300  
 80 TTTAATAATA TCTACAGGGA ATTACAGGAA GAAGAGTCTC TGGTGGAAAC GTTTGTGTTT 3360  
 GAAGCCTTGG TGATATACAT GGAGAGTCTG GCCTTAGCAC ATGCAGATGA GAAGTCTTAA 3420  
 GGTACAATTC AACAGTGTG TGATGCCATT GATCACCTAT GCGCATCAT TGAAGAAG 3480  
 CATGTTTCTT TAAATAAAGC AAAGAAACGA CGTTTGGCCG GAGGATTTCC ACCTTCGCA 3540  
 TCAATGTGTT TATGATCTT GGTCAAGTGG CTTTATGCTC ATTTGGGAG GCCCAGACA 3600  
 85 GAATGTGAC ACAATTCATT TGAACCTTT TATAAATTG TCCTTTATT GCCAGGCAAC 3660  
 AGATCCCTTA ATTTGTGGCT GAAAGATGTT CTCAAGGAAG AAGGTGCTC TTTTCTCATC 3720  
 AACACCTTTG AGGGGGTGG CTGTGGCCAG CCTCGGGCA TCCTGGCCA GCCACCTCTC 3780  
 TTGTACCTTC GGGGGCCATT CAGCCTGCAG GCCACGCTAT GCTGGCTGGA CTTGCTCTG 3840



226

	TACAACCCACC	TTGCTGAGTG	GAAATCACTT	GAATACTGTT	CTACAGCCAG	TATAGACAGT	9120
	GAGAACCCCC	CAGACTTAAA	TAAAATCTGG	AGTGAACCAT	TTTATCAGGA	AACATATCTA	9180
	CCTTACATGA	TCCGACAGCA	GCTGAAGCTG	CTGCTCCAGG	GAGAGGCTGA	CCAGTCCCTG	9240
5	CTGACATTTA	TTGACAAAGC	TATGCACGGG	GAGCTCCAGA	AGGCGATTCT	AGAGCTTCAT	9300
	TACAGTCAAG	AGCTGAGTCT	GCTTTACCTC	CTGCAAGATG	ATGTTGACAG	AGCCAAATAT	9360
	TACATTCAAA	ATGGCATTCA	GAGTTTTATG	CAGAATTATT	CTAGTATTGA	TGTCCTCTTA	9420
	CACCAAGATA	GACTCAACAA	ATTGACAGTCT	GTACAGGCTT	TAACAGAAAT	TCAGGAGTTC	9480
	ATCAGCTTTA	TAAGCAAAAC	AGGCAATTTA	TCATCTCAAG	TTCCCTCTAA	GAGACTTCTG	9540
10	AACACCTGGA	CAAAACAGATA	TCCAGATGCT	AAAATGGACC	CAATGAACAT	CTGGGATGAC	9600
	ATCATCAAA	ATCGATGTTT	CTTTCTCAGC	AAAATAGAGG	AGAAGCTTAC	CCCTCTTCCA	9660
	GAAGATAATA	GTATGAATGT	GGATCAAGAT	GGAGACCCCA	GTGACAGGAT	GGAGGTGCAA	9720
	GAGCAGGAAG	AAGATATCAG	CTCCCTGATC	AGGAGTTGCA	AGTTTTCAT	GAAAATGAAG	9780
	ATGATAGACA	GTGCCCGGAA	GCAGAACAA	TTCTCACTTG	CTATGAACT	ACTGAAGGAG	9840
15	CTGATAAAG	AGTCAAAAC	CAGAGACGAT	TGGCTGGTGA	GCTGGGTGCA	GAGCTACTGC	9900
	CGCTGAGCC	ACTGCCGAG	CCGGTCCAG	GGCTGCTCTG	AGCAGGTGCT	CACCTGTGCTG	9960
	AAAACAGTCT	CTTGTGTTGA	TGAGAACCAAC	GTGTCAAGCT	ACTTAAGCA	AAATATTCTG	10020
	GCTTTCCGTC	ACCAGAACAT	TCTCTTGGGT	ACAACCTTAC	GGATCATAGC	GAATGCTCTC	10080
	AGCAGTGAGC	CAGCTTGCC	TGCTGAAATC	GAGGAGGACA	AGGCTAGAAG	AACTCTTAGAG	10140
20	CTTTCTGGAT	CAAGTTCAGA	GGATTGAGAG	AAGGTGATCG	CGGGTCTGTA	CCAGAGAGCA	10200
	TTCCAGCACC	TCTCTGAGCG	TGTGACGGCG	GCTGAGGAGG	AGGCCACGCC	TCCTCTCTGG	10260
	AGCTGTGGCG	CTGCGAGCTG	GGTGATTGAT	GCTTACATGA	CGCTGGCAGA	TTTCTGTGAC	10320
	CAACAGCTGC	GCAAGGAGGA	AGAGAATGCA	TCAGTTATTG	ATTCTGCAGA	ACTGCAGGCG	10380
	TATCCAGCAC	TTGTGTGGA	GAAAATGTTG	AAAGCTTTAA	AAATAAATTC	CAATGAAGCC	10440
25	AGATTGAAGT	TTCTAGATT	ACTTCAGATT	ATAGAACCGT	ATCCAGAGGA	GACTTTGAGC	10500
	CTCATGACAA	AAGAGATCTC	TTCCGTTCCT	TGCTGGCAGT	TCATCAGCTG	GATCAGCCAC	10560
	ATGCTGGCCT	TACTGGACAA	AGACCAAGCC	GTGCTGTTC	AGCACTCTGT	GGAGAAATC	10620
	ACTGATAACT	ACCCGAGGC	TATTGTTTAT	CCCTTCATCA	TAAGCAGCGA	AAGCTATTCC	10680
	TTCAAGGATA	CTTCTACTGG	TCATAAGAAT	AAGGAGTTTG	TGGCAAGGAT	TAAAAGTAAG	10740
30	TTGGATCAAG	TTCCTAGATT	TCAGATTTT	ATTAATGCTT	TAGATCAGCT	CTCTAATCCT	10800
	GAATCTGCTC	TTAAGGATTG	GAGCAATGAT	GTAAGAGCTG	AACTAGCAAA	AAACCCCTGT	10860
	AATAAAAAAA	AACTTGAAAA	AATGTATGAA	AGAATGTATG	CAGCCTTGGG	TGACCCAAAG	10920
	GCTCCAGGCC	TGGGGGCTTT	TAGAAGGAAG	TTTATTGAGA	CTTTTGGAAG	AGAATTGTAT	10980
	AAACATTITG	GGAAAGGAGG	TTCTAAACTA	CTGAGAATGA	AGCTCAGTGA	CTTCAACGAC	11040
35	ATTACCAACA	TGCTACTTTT	AAAAATGAAC	AAAGACTCAA	AGCCCCCTGG	GAATCTGAAA	11100
	GAATGTTTAC	CCTGGATGAG	CGACTTCAAA	GTGGAGTTCC	TGAGAAATGA	GCTGGAGATT	11160
	CCCGTTCAGT	ATGACGTTAG	GGGAAAGCCA	TTGCCAGAGT	ACCAAGTGGG	AATCGCCGGG	11220
	TTTGATGAGC	GGGTGACAGT	CATGGCGTCT	CTGCGAAGGC	CCAAGCGCAT	CATCATCCGT	11280
40	GGCCATGAGC	AGAGGGAACA	CCCTTTCCTG	GTGAAGGGTG	GCGAGGACCT	CGCGCAGGAC	11340
	CAGCGCGTGG	AGCAGCTCTT	CCAGGTCTATG	AATGGGATCC	TGGCCCAAGA	CTCCGCTGTC	11400
	AGCCAGAGGG	CCCTGCGAGT	GAGGACCTAT	AGCGTTGTGC	CCATGACCTC	CAGGTTAGGA	11460
45	TTAATTGAGT	GGCTTGAAAA	TACTGTTACC	TTGAAGGACC	TTCTTTTGAA	CACCATGTCC	11520
	CAAGAGGAGA	AGGCGGCTTA	CCTGAGTGAT	CCCAAGGCAC	CGCCGTGTGA	ATATAAAGAT	11580
	TGGCTGACAA	AAATGTGAGG	AAACATGAT	GTGGAGCTTT	ACATGCTAAT	GTATAAGGGC	11640
	GCTAATCGTA	CTGAACAGT	CACGTCTTTT	AGAAAACGAG	AAAGTAAAGT	GCCTGCTGAT	11700
	CTCTTAAAGC	GGGCTTCTGT	GAGGATGAGT	ACAAGCCCTG	AGGCTTTCTT	GGCGCTCCGC	11760
50	TCCCACTTGC	CCAGCTCTCA	CGCTCTGATA	TGCTACAGCC	ATCGGATCCT	CGGGATTGGA	11820
	GACAGACATC	TGAACAACCT	TATGGTGCC	ATGGAGACTG	GCGCGTGAT	CGGGATGAGC	11880
	TTTGGGATG	CGTTTGGATC	CGCTACACAG	TTCTGCGCAG	TCCTGAGTT	GATGCCTTTT	11940
	CGGCTAACTC	GCCAGTTTAT	CAATCTGATG	TTACCAATGA	AAGAAACGGG	CCTTATGTAC	12000
	AGCATCATGG	TACACGCACT	CCGGGCTTTC	CGCTCAGACC	CTGGCCTGCT	CACCAACACC	12060
	ATGGATGTGT	TTGTCAAGGA	GCCTCTCTTT	GATTGGAAGA	ATTTTGAACA	GAAAATGCTG	12120
	AAAAAAGGAG	GGTCATGGAT	TCAAGAAATA	AATGTTGCTG	AAAAAATTG	GTACCCCGGA	12180
55	CAGAAATAT	GTACCGTTAA	GAGAAAGTTA	GCAGGTGCCA	ATCCAGCAGT	CATTACTTGT	12240
	GATGAGCTAC	TCTTGGGTCA	TGAGAAAGCC	CCTGCTTCA	GAGACTATGT	GGCTGTGGCA	12300
	CGAGGAAGCA	AAGATCACAA	CATTGTGCC	CAAGAACCCG	AGAGTGGGCT	TTCAGAAGAG	12360
	ACTCAAGTGA	AGTGCTCTAT	GGACCAGGCA	ACAGACCCCA	ACATCCTTGG	CAGAACCTGG	12420
	GAAGGATGGG	AGCCCTGGAT	GTGAGGCTG	TGGGAGTCTG	CAGATAGAAA	GCATTACATT	12480
	GTTTAAAGAA	ATCTTATATC	TTTGGTTGGC	AGCATTCCAT	GAGCTGATTT	TCTTGAAACA	12540
60	CTAAAGAGAA	ATGCTTTTGT	TGCTACAGTT	TGCTAGCATG	AGTTTAAATC	AAGATTATGA	12600
	TGAGTAAATG	TGATAGGTTT	AAATCAAGA	TAAGGTTATA	GTAACATCAA	AGATTAGGTG	12660
	AGGTTTATAG	AAAGATAGAT	ATCCAGGCTT	ACCAAAGTAT	TAAGTCAAGA	ATATAATATG	12720
	TGATCAGCTT	TCAAGCATT	TACAAGTGCT	GCAAGTTAGT	GAAACAGCTG	TCTCGTAA	12780
	TGGAGGAAT	TGGGGGAAGC	CTTGAATGCT	CCTTCTGGTT	CTGGCAGATT	GGAAAGCACA	12840
65	CTCAGAAAGC	TTTCATCACA	AGATTTTGGG	AGAGTAAAGC	TAAGTATAGT	TGATGTAAAC	12900
	TTGTAGAAGC	AGCATAGGAA	CAATAAGAAC	AATAGGTAAA	GCTATAATTA	TGGCTTATAT	12960
	TTAGAAATGA	CTGCATTGGA	TATTTTAGGA	TATTTTCTTA	GGTTTTTCTC	TTTCAATTTA	13020
	TTCTCTTCTA	GTTTTGACAT	TTTATGATAG	ATTTGCTCTC	TAGAAGGAAA	CGTCTTTATT	13080
	TAGGAGGGCA	AAAATTTTGG	TCATAGCATT	CACTTTGTCT	ATTCCAATCT	ACAACTGGAA	13140
70	GATACATAAA	AGTGCTTTGC	ATTGAATTTG	GGATAACTTC	AAAAATCCCA	TGGTTGTTGT	13200
	TAGGATAGAT	ACTAAGCATT	TCAGTTCCAG	GAGAATAAAA	GAAATTCCTA	TTTGAAATGA	13260
	ATTCTCTATT	TGGAGGAAAA	AAAGCATGCA	TTCTAGCACA	ACAAGATGAA	ATTATGGAAT	13320
	ACAAAAGTGG	CTCCTTCCCA	TGTGCACTCC	CTGTCCCTCC	CGGCCAGTCC	TCCACACCCA	13380
	AACTGTTTCT	GATTGGCTTT	TAGCTTTTGT	TTGTTTTTTT	TTTTCTTCTT	AACACTTGTA	13440
75	TTTGAAGGCT	CTTCTGTGAT	TTTGAGAAGT	ATACTCTTGA	GTGTTTAAAT	AAGTTTTTTT	13500
	CCAAAAGTA						

Seq ID NO: 99 Protein sequence:  
Protein Accession #: NP\_008835.5

80	1	11	21	31	41	51	
	MAGSGAGVRC	SLRLRLQETLS	AADRCGAALA	GHLRLRLGQ	ECVLSSSPAV	LALQTSLVFS	60
	RDFGLLVFVR	KSINSLIPRE	CREELKFLC	IFLEKMGQKI	APYSVEIKNT	CTSVYTKDRA	120
85	AKCKIPALDL	LILKILQTPRS	SRLMDEPKIG	ELFSKPYGEL	ALKKKIPDTV	LEKVVYELLGL	180
	LGEVHPSEMI	NNAENLFPAP	LGELKTQMTS	AVREPKLPLVL	AGCLKGLSSL	LCNFTKSMEE	240
	DPQTSREIFN	FVLKAIKRPQI	DLKRYAVPSA	GLRLFPALHAS	QFSTCLLDNY	VSLFEVLLKW	300
	CAHTNVELKK	AALSALBSFL	KQVSNMVAKN	AEMHKNKLQY	FMEQFYGIIR	NVDSNNKELS	360

	IAIRGYGLFA	GPCKVINAKD	VDFMYVELIQ	RCKQMLTQT	DTGDDRVYQM	PSFLQSVASV	420
	LLYLDTVPEV	YTPVLEHLVV	MQIDSFPQYS	PKMQLVCCRA	IVKVFLALAA	KGPFVLRNCIS	480
	TVVHQLTRI	CSKPVVLPKG	PESESEDHRA	SGEVRTGKWK	VPTYKYVDL	FRHLLSSDQM	540
5	MDSILADEAF	FSVNSSSSSL	NHLLYDEPVK	SVLKIVKELD	LTLBIQTVGE	QBNGDEAPGV	600
	WMIPTSDPAA	NLHPAKPKDF	SAPINLVEFC	REILPEKQAE	FFPEWVVSFS	YELLQSTRL	660
	FLISGFYKLL	SITVRNAKKI	KYFEGVSPKS	LKHSPEDEPK	YSCFALFVKF	GKEVAVVMKQ	720
	YKDELLASCL	TFLLSLFENI	IELDVRAVVP	ALQMAFKLGL	SYTPLABVGL	NALEWSIYI	780
	DRHVMQPYTK	DILPCLODYL	RTSALSDETK	NNWEVSALSR	AAQGFNKKV	LKHLKKTKNL	840
10	SSNEAISLEE	IRIRVVQMLG	SLGGQINKNL	LTVTSSEDM	KSYVANDREK	RLSFVVPFRE	900
	MKPVIFLDVF	LRPVTELALT	ASDRQTKVAA	CELLHSMVMF	MLGKATQMP	GGQAPPMYQ	960
	LYKRTFPVLL	RLACDQVQVT	RQLYEPLVMQ	LIHWFTNNKK	FRSQDTVALL	EAILDGIQDP	1020
	VDSTLRDFCG	RCIREFLKWS	IKQITPQQQE	KSPVNTKSLP	KRLYSLALHP	NAPKRLGASL	1080
	APNNIYREFR	EEBSLVEQFV	FEALVIYMES	LALAHADERS	LGTIQCCDA	IDHLCRITIEK	1140
15	KHVSINKAKK	RRLPRGPPPS	ASLCLDLVK	WLLAHCGRPQ	TECRHKSIEL	FYKFPVLLPG	1200
	NRSPLNLKAK	VLKEEGVSFL	INTFEGGCGG	QPSGILAQPT	LLYLGRPFSL	QATLCWLDLL	1260
	LAALECYNTF	IGERTVGAQL	VLGTEAQSSL	LKAVAFPLES	IAMHDIIAAE	KCFGTAAGN	1320
	RTSPQSGERY	NYSKCTVVVR	IMEFTTLLN	TSPEGWKLK	KDLCNTHLMR	VIVQTLCEPA	1380
	SIGFNIQDVQ	VMALPDPVCV	NLMKALKMSP	YKDIETHLR	EKITAQSIIE	LCAVNLVQPD	1440
20	AQVDRSRLAA	VVSACQQLER	AGLLENILPS	QSTDLLHSVG	TELLSLVYKQ	IAPGDERQCL	1500
	PSLDLSCQQL	ASGLLELAPA	FGGLCERLVS	LLNLPAVLST	ASLGSSQGSV	IHFSGHEYPY	1560
	SLFSETINLE	LLAKNLDLAV	ELMQSSVDNT	KMVSANLNGM	LQSQFRERAN	KHQGLKGLAT	1620
	TILQHWKCCD	SNWAKDSPLE	TKMAVLALLA	KILQIDSSVS	FNTSHGSPFE	VFTTYISLLA	1680
	DTKLDLHLKG	QAVTLPLPFT	SLTGGSLLEL	RRVLEQLIVA	HFPMQSREPP	PGTFRPNVY	1740
25	DCMKFPLDAL	ELSQSPMLLE	LMTEVLCREQ	QHVMEELFQS	SFRRIARRGS	CVTQVGLLES	1800
	VYEMFRKDDP	RLSFTQSFV	DRSLTLTLLH	CSLDALREFF	STIVVDAIDV	LKSFTKLNE	1860
	STFDQITKK	MGYKILDDVM	YSRLPKDDVH	AKESKINQVP	HGSCITEGNE	LTKTLIKLCY	1920
	DAFTENMAGE	NQLLERRRLY	HCAAYNCAIS	VICCVFNEKL	FYQGFLESEK	PEKNLLIFEN	1980
	LIDLKRRYNF	PVEVEVPMER	KKKYIEIRKE	AREAANGSD	GPSYMSLSLY	LADSTLSEEM	2040
30	SQDPSTGVQ	SYSYSSQDPR	PATGRPRRR	QRDPTVHDDV	LELEMDLNR	HECMAPLAL	2100
	VKMHRSGLP	PGGEEDSVPR	DLPWMKFLH	GKLGNIPIVL	NIRLEFLAKV	INTEEVFRPY	2160
	AKHNLSPILQ	LAASNNNGE	GIHYMVVEIV	ATILSWTGLA	TPTGVPKDEV	LANRLNLFNM	2220
	KHVFHPKRAV	FRHNLBIKT	LVECNKDCLS	IPYRLIFEKF	SGKDPNSKDN	SVGIQLGLIV	2280
35	MANDLPYPDP	QCGIQSSEYP	QALVNMSPV	RYKEVYAAAA	EVLGLILRVY	MERONILEES	2340
	LCELVARQLK	QHQNTMEDKF	IVCLANKVKS	FPPLADRFMN	AVFFLLPKPH	GVLTLCLEV	2400
	LVLCRVGMTS	LYFQLKSKDF	VQVMREHDD	RQKVCLDIY	KMPKPKPVE	LRELLNPVVE	2460
	FVSHPTTCR	BOMYNILMWI	HDNYRDPES	TDNDSQEIFK	LAKDVLIOGL	IDENPGLQLI	2520
	IRNFWSHETR	LPSNTLDRLL	ALNSLYSPKI	EVHFLSLATN	FLLEMTSMSP	DYPNPMFEHP	2580
	LSECEPQETT	IDSDWRFRST	VLTPMPVETQ	ASQGTQLTQT	QEGSLSARWR	VAGQIRATQO	2640
40	QHDFTLTQTA	DGRSSFDWLT	GSSTDPLVDH	TSPSSDLSLF	AHKRSERLQR	APLKSVPQDF	2700
	GKRLGLPGD	EVNDKVGAA	GRTDLLRLRR	RFMRDQEKLS	LMYARKGVAE	QKREKEIKSE	2760
	LKMKDAQVQ	LYRSYRHGDL	PDIQIKHSSL	ITPLQAVAR	DPIIAKQLFS	SLFSGILKEM	2820
	DKFKTLSEKN	NITQRLQDQF	NRFPLNTTFS	FPFVSCIQD	ISQQAALLS	LDPAAVSAGC	2880
	LASLQPPVGI	RLLEBALLRI	LPAELPAKRV	RGKARLPFDV	LRWVELAKLY	RSIGEVDVLR	2940
45	GIFTSEIGTK	QITQALLAE	ARSDYSEAAK	QYDEALNKQD	WVDGEPTAE	KDFWELASLD	3000
	CYNHAEWKS	LEYCASTASID	SENPPDLNKI	WSEPFYQETY	LPYMIKSKLK	LLLQGGADQS	3060
	LLTFIDKAMH	GELQKALLEL	HYSQELSLLY	LLQDDVDRAK	YYIQNGIQSF	MQNYSSIDVL	3120
	LHQSLRLTKI	SVQALTEIQE	FISFISKQGN	LSSQVPLKRL	LNTWTNRYPD	AKMDPMNIWD	3180
50	DIITNRCEFL	SKIEBKLTPL	PEDNSMNVQD	DGDPSDRMEV	QEQEEDISSL	IRSCPKSMKM	3240
	KMIDSARRQN	NPSLAMKLLK	ELHKEKSTRD	DNLVSWVQSY	CRLSHCRSRS	QGCSEQVLTV	3300
	LKTVSLLDEN	NVSSYLSKNI	LAFRDQNIL	GTYYRIIANA	LSSEPACLAE	IEEDKARRIL	3360
	ELSGSSSEDS	EXVIAAGLYR	AFQHLSEAVQ	AAEKEAQQPS	WSCGPAAGVI	DAYMTLADFC	3420
	QDQLKKEEN	ASVIDSAELQ	AYPALVVERM	LKALKNSNE	ARLKFPRLQV	IERYPEETL	3480
	SLMTKEISGV	PCWGTISNIG	HMVALLDKQ	AVAVQHSVEE	ITDNYQOAVI	YPFIISSESY	3540
55	SPKDTSTGKH	NKEFVARIKS	KLDQGGVIQD	FINALDQLSN	PELLFKDWSN	DVRAELAKTP	3600
	VNKKNIERMY	ERMVYALGDP	KAPGLGAPRR	KFIQTGKEF	DKHFGKGGSK	LLRMKLSDPN	3660
	DTTMMLLML	NKDSKPPGNL	KECSPWMSDF	KVEFLRNELE	IPGQYDGRGK	PLPEYHVRIA	3720
	GFDERVTYMA	SLRRFKRIII	RGHDEREHPP	LVKGGEDLRQ	DQVBEQLPQV	MNGILAQDSA	3780
	CSQALQRLT	YSVFMPTSR	GLIENLENTV	TLKDLLLNTM	SQEEKAAYLS	DPRAPPCEYK	3840
60	DNLTKMSGKH	DVGAYMLMYK	GANRTETVTS	FRKRESKVP	DLLKRAFVRM	STSPAPLAL	3900
	RSHFASSHAL	ICISHWILGI	GDRHLNFMV	AMETCGVIGI	DPGHAFQSAT	QFLPVPELMP	3960
	FRLTRQFINL	MLPMEKGTGLM	YSIMVHALRA	FRSDPGLLTN	TMDVVFKEPS	FDWKNFEQKM	4020
	LKKGGSWIQE	INVAEKWNPY	RQKICYAKRK	LAGANPAVIT	COELLIGHEK	APAFRDYVAV	4080
	ARGSDENIR	AQEPESGLSE	ETQVKCLMDQ	ATDPNILGRT	WEGWEPWM		

Seq ID NO: 100 DNA sequence  
Nucleic Acid Accession #: NM\_000673  
Coding sequence: 101-1225

	1	11	21	31	41	51	
70	ATGTGAAGGC	ACAAGCTGCT	GTTATATACA	ACAGAGTGAA	CTGAGCATCA	GTGAGAAAA	60
	GTCTATGTTT	GCAGAAATAC	AGATCCAAGA	CAAGACAGS	ATGGGCACTG	CTGGAAAAGT	120
	TATTAAATGC	AAAGCAGCTG	TGCTTTGGGA	GCAGAAAGCA	CCCTTCTCCA	TTGAGGAAAT	180
75	AGAAGTGTCC	CCACCAAGA	CTAAGAAGT	TGCGATTAG	ATTTTGGCCA	CAGGAATCTG	240
	TGCGACAGAT	GACCATGTGA	TAAAAGGAAC	AATGCTGTCC	AAGTTTCCAG	TGATTGTGGG	300
	ACATGAGGCA	ACTGGGATTG	TAGAGAGCAT	TGAGAGAGCA	GTGACTACAG	TGAAACCAGG	360
	TGACAAAGTC	ATTCCTCTCT	TTCTGCCACA	ATGTAGAGAA	TGCAATGCTT	GTGCGAACCC	420
	AGATGGCAAC	TTCTGATTAT	GGAGCGATAT	TACTGGTCGT	GGAGTACTGG	CTGATGGCAC	480
80	CACCAAGATT	ACATGCAAGG	GCAAAACAGT	ACACCATTCT	ATGAACACCA	GTACATTTC	540
	CGAGTACACA	GTGGTGGATG	AATCTTCTGT	TGCTAAGATT	GATGATGCAG	CTCCTCCTGA	600
	GAAAGTCTGT	TTAATTGGCT	GTGGGTTTTC	CACTGATAT	GGCGCTGCTG	TTAAACTCGG	660
	CAAGGTCAAA	CCTGGTTCCT	CTTGGCTCGT	CTTTGGCCTG	GGAGGAGTGT	GCCTGTCTAGT	720
	CATCATGGGC	TGTAAGTCAG	CTGGTGATCT	TAGGATCATT	GGGATTGACC	TCAACAAAGA	780
85	CAAAATTGAG	AAGGCCATGG	CTGTAGGTGC	CACTGAGTGT	ATCAGTCCCA	AGGACTCTAC	840
	CAAAACCATC	AGTGAGGTGC	TGTAGAAAT	GACAGGCAAC	AACGTGGGAT	ACACCTTTGA	900
	AGTTATTGGG	CATCTTGAA	CCATGATTGA	TGCCCTGGCA	TCCTGCCACA	TGAACATATG	960
	GACCAAGCGT	GTTGTAGGAG	TTCCTCCATC	AGCCAGATG	CTCACCTATG	ACCGATGTT	1020

GCTCTTCACT GGACGCACAT GGAAGGGATG TGTCTTGGGA GGTTTGAAAA GCAGAGATGA 1080  
 TGTCCCAAAA CTAGTGACTG AGTTCCCTGGC AAAGAAATTT GACCTGGACC AGTTGATAAC 1140  
 TCATGTTTAA CCATTAAAAA AAATCAGTGA AGGATTTGAG CTGCTCAATT CAGGACAAAG 1200  
 CATTCGAACG GTCCTGACGT TTTGAGATCC AAAGTGGCAG GAGGCTGTGT TTGTCATGGT 1260  
 GAACGTGAGT TTCTCTGTGT AGAGTTCCTT CATCTGAAAT CATGTATCTG TCTCACAAT 1320  
 ACAAGCATAA GTAGAAGATT TGTGAAGAC ATAGAACCCT TATAAAGAAT TATTAACCTT 1380  
 TATAAACATT TAAGTCTCTG TGAGCACCTG GGAATTAGTA TAATAACAAT GTTAATATTT 1440  
 TTGATTTACA TTTTGTAAAG CTATAATTGT ATCTTTTAAG AAAACATACA CTTGGATTTT 1500  
 TATGTTGAAA TGGAGATTTT TAAGAGTTTT AACCACTGTC TGCAGATATA TAACTCAAAA 1560  
 CAGATATAGC GTATAAAGAT ATAGTAAATG CATCTCCAG AGTAATATTC ACTTAACACA 1620  
 TTGAAACTAT TATTTTTTAG ATTTGAATAT AAATGTATTT TTTAAACACT TGTATGAGT 1680  
 TAACCTGGAT TACATTTTGA AATCAGTICA TTCCATGATG CATATTACTG GATTAGATTA 1740  
 AGAAGACAG AAAAGATTAA GGGACGGGCA CATTTTTCAG CGATTAAAGT TCATCATTAC 1800  
 ATAACTTGGT GAAACTGAAA AAGTATATCA TATGGGTACA CAAGGCTATT TGCCAGCATA 1860  
 TATTAATATT TTAGAAATA TTCTTTTGT AATACTGAAT ATAAACATAG AGCTAGAGTC 1920  
 ATATTATCAT ACTTATCATA ATGTTCAATT TGATACAGTA GAATTGCAAG TCCCTAAGTC 1980  
 CCTATTCACT GTGCTTAGTA GTGACTCCAT TTAATAAAAA GTGTTTTTAG TTTTAAACAA 2040  
 CTAACCG

Seq ID NO: 101 Protein sequence:  
 Protein Accession #: NP\_000664

1 11 21 31 41 51  
 MGTAGKVIK KAAVLWEQKQ PFSIEIEIVA PPKTKEVRIK ILATGICRTD DHVIKGTMVS 60  
 KFPVIVGHEA TGIIVSIGEG VTTVKPGDKV IPLFLPQCRE CNAACRNPEDGN LCIRSDITGR 120  
 GVLADGTRF TCKGKPVHFF MNTSTFTETV VVDESSVAKI DDAAPEKVC LIGCFSTGY 180  
 GAAVKTGKVK PGSTCVVFGI GGVGLSVIMG CKSAGASRII GIDLKDKFE KAMAVGATEC 240  
 ISPKDSKPI SEVLSEMTGN NVGYTFEVIQ HLETMIDALA SCHMNYGTSV VVGVPSSAKM 300  
 LTYDPMILLFT GRTWKGCVFQ GLKSRDDVPK LVTEFLAKKF DLDQLITHVL PFKKISEGFE 360  
 LINSQQSIRT VLTFF

Seq ID NO: 102 DNA sequence  
 Nucleic Acid Accession #: NM\_006783.1  
 Coding sequence: 1..786

1 11 21 31 41 51  
 ATGGATTGGG GGACGCTGCA CACTTTTCATC GGGGGTGTCA ACAACACTC CACCAGCATC 60  
 GGGAAGGTGT GGATCAGAGT CATCTTTATT TTCGAGTCA TGATCCTAGT GGTGGCTGCC 120  
 CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC 180  
 AAAAATGTGT GCTATGACCA CTTTTCCTCC GTGTCCCACT TCCGGCTGTG GGCCTCCAG 240  
 CTGATCTTGT TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCTTA CTACAGGCAC 300  
 GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360  
 ATTAAGAAAG ACAAGGTTCG GATAGAGGGG TCGCTGTGGT GGAGCTACAC CAGCAGCATC 420  
 TTTTTCGAAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACAATGGG 480  
 TACCACCTGC CCTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGTACTGC 540  
 TTTATTTCTA GGCCCAACAGA GAAGACCGTG TTTACCAATT TTATGATTTC TGCCTCTGTG 600  
 ATTTGCTATG TGCTTAAAGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTTAGG 660  
 AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720  
 CAGAATGAAA TGAATGAGCT GATTTTCAGAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780  
 AGCTAA

Seq ID NO: 103 Protein sequence:  
 Protein Accession #: NP\_006774.1

1 11 21 31 41 51  
 MDWGLHTFI GGVNKHSTSI GKVWITVIFI FRVMILVVAA QEVNGDEQED FVCNTLQPGC 60  
 KNVCYDHFFP VSHIRLWALQ LIPVSTPALL VAMHVAYYRH RTTRKFRERGE KRNDPDIED 120  
 IKKHVRIEG SLNWTYSSSI PFRIIFEAAP MYVPYFLYNG YHLPWVLKCG IDPCPNLVDG 180  
 FISRPTEKTV FTIFMISASV ICMLLNVAEL CYLLKVCFR RSKRAQTQKN HPNHALKESK 240  
 QNEMNELISD SQQNAITGFP S

Seq ID NO: 104 DNA sequence  
 Nucleic Acid Accession #: NM\_020411  
 Coding sequence: 86-526

1 11 21 31 41 51  
 GGACCTGGGA AGGAGCATAG GACAGGGCAA GCGGGGATAA GGAGGGGCAC CACAGCCCTT 60  
 AAGGCACGAG GGAACCTCAC TGGCGATGCT CCTTTGGTGC CCACCTCAGT GCGCATGTTT 120  
 ACTGGGCGTC TTCCCATCGG CCCCTTCGCC AGTGTGGGGA ACGCGGCGGA GCTGTGAGCC 180  
 GCGGACTCGG GTCCCTGAGG TCTGGATTCT TTCTCCGCTA CTGAGACAG GCGGACACAC 240  
 ACAACACAG AACCAACAG CCAGTCCCAG GAGCCAGTA ATGGAGAGCC CCAAAAAGAA 300  
 GAACACGAG CTGAAGTTCG GGATCCCTACA CCTGGGCGAG AGACAGAAGA AGATCAGGAT 360  
 ACAGCTGAGA TCCAGTGGG CGACATGGAA GGTGATCTGC AAGAGCTGCA TCAGTCAAC 420  
 ACCGGGGATA AATCTGGAAT TGGGTTCCGG CGTCAAGGTG AAGATAATAC CTAAGAGGA 480  
 ACACGTGAAA ATGCCAGAAG CAGGTGAAGA GCAACACAA GTTTAAATGA AGACAAGCTG 540  
 AAACACGCA AGCTGGTTTT ATATTAGATA TTTGACTTAA ACTATCTCAA TAAAGTTTTG 600  
 CAGCTTTCAC CAAAAAATA AAAAAA

Seq ID NO: 105 Protein sequence:  
 Protein Accession #: NP\_065144.1

1 11 21 31 41 51

MLLWCPFQCA CSLGVFPFAP SPVWGTTRSC EPATRVPEVM ILSPLL RHGG HTQTQNHSTAS 60  
 PRSPVMESPK KKNQQLKVG I LHLGSRQKKI RIQLRSQCAT WKVICKSCIS QTPGINLDLG 120  
 SGVXVKIIPK EEHCMPFAG EEQFQV

Seq ID NO: 106 DNA sequence  
 Nucleic Acid Accession #: J04129  
 Coding sequence: 99-587

1 11 21 31 41 51  
 CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCTCCTGC 60  
 TCACCTCTGG GGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120  
 AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CCTGGCACTC CATGGCCATG GCGACCAACA 180  
 ACATCTCCCT CATGGCGACA CTGAAGGCCC CTCTGAGGGT CCACATCAAC TCACTGTGTC 240  
 CCACCCCCGA GGACAACTTG GAGATCGTTC TGACACAGATG GGAGAACAAC AGCTGTGTGT 300  
 AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATAACG 360  
 TGGCGAAGCA GGCACCGCTG CTCGATACTG ACTACGACAA TTTCTGTGTT CTCTGCCTAC 420  
 AGGACACCCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCTCTGTGG 480  
 AGGACATGAT GATCATGCGA GGATTCATCA GGGCTTTCAG GCCCCTGCCC AGGCACCTAT 540  
 GGTACTTGCT GGACTTGAAA CAGATGGAAG AGCCGTGCCG TTTCTAGCTC ACCTCGGCCT 600  
 CCAGGAAGAC CAGACTCCCA CCCTCCACA CCTCCAGAGC AGTGGGACTT CCTCTGCCC 660  
 TTTCAAAGAA TAACCAACAG TCAGAAGACG ATGACGTGGT CATCTGTGTC GCCATCCCT 720  
 TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCCTGGGGGC AGAGTCTCTG 780  
 GCAGAGGTTA TTAATAAACC CTTGGAGCAT G

Seq ID NO: 107 Protein sequence:  
 Protein Accession #: AAA60147

1 11 21 31 41 51  
 MDIPQTKQDL ELPKLAGTWH SMAMATNNIS LMATLKAPLR VHITSLLPTR EDNLEIVLHR 60  
 WENNSCIVEKK VLGEKTKGNPK KFKINYTVAN EATLLD TDYD NFLPLCLQDT TTPIQSMMQ 120  
 YLARVLVEED EIMGGPIRAF RPLPRHLWYL LDLQMEBPC RF

Seq ID NO: 108 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 48-794

1 11 21 31 41 51  
 TCCAGGCAG CAGITAGCCC GCCGCCCGCC TGTGTGTCCC CAGAGCCATG GAGAGAGCCA 60  
 GTCTGATCCA GAAGGCCAAG CTGGCAGAGC AGGCGAAGC CTATGAGGAC ATGGCAGCCT 120  
 TCATGAAAGG CGCCGTGGAG AAGGGCGAGG AGCTCTCCTG CGAAGAGCGA AACCTGCTCT 180  
 CAGTAGCCTA TAAGAACGTG GTGGCGGGCC AGAGGGCTGC CTGGAGGGTG CTGTCCAGTA 240  
 TTGAGCAGAA AAGCAACGAG GAGGGCTCGG AGGAGAAGGG GCCCGAGGGT CGTGAGTACC 300  
 GGGAGAAGGT GGAGACTGAG CTCAGGGCGG TGTGCGACAC CGTGTGGGCG CTGTGGACA 360  
 GCCACCTCAT CAAGGAGGCC GGGGAGCGCG AGAGCGGGT CTCTACTCTG AAGATGAAGG 420  
 GTGACTACTA CCGCTACCTG GCCGAGGTGG CCACGGGTGA CGACAAGAAG CGCATCATTG 480  
 ACTCAGCCCG GTACGCTTAC CAGGAGGCCA TGGACATCAG CAAGAAGGAG ATGCCGCCCA 540  
 CCAACCCCAT COGCTGGGCG CTGGCCCTGA ACTTTTCGGT CTCTCACTAC GAGATCGCCA 600  
 ACAGCCCCGA GGAGGCCATC TCTCTGGCCA AGACCACTTT CGAGAGGGCC ATGGCTGATC 660  
 TGCACACCTC CAGCGAGGAC TCCTACAAAG ACAGCACCTT CATCATGCAG CTGCTGGAG 720  
 ACAACCTGAC ACTGTGAGCG GCCGACACG CCGGGGAAGA GGGGGGCGAG GCTCCCGAG 780  
 AGCCCCAGAG CTGAGTGTG CCCGCCACCG CCCCGCCCTG CCCCTCCAG TCCCCACCC 840  
 TGCCGAGAGG ACTAGTATGG GGTGGGAGGC CCCACCTTC TCCCTTAGGC GCTGTCTCTG 900  
 CTCCAAAGGG CTCGTGGAG AGGGACTGGC AGAGCTGAGG CCACCTGGGG CTGGGGATCC 960  
 CACTCTTCTT GCAGCTGTG AGGCACCTA ACCACTGGTC ATGCCCCAC CCCTGCTCTC 1020  
 CGCACCCGCT TCCTCCGAC CCCAGGACCA GGCTACTTCT CCCCTCCTCT TGCTCCCTC 1080  
 CTGCCCTGCG TGCTCTGAT CGTAGGAATT GAGGAGTGT CCGCCTTGTG GCTGAGAACT 1140  
 GGACAGTGGC AGGGGCTGGA GATGGGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG 1200  
 CGCGCGCGCC AGTGCAGAC CGAGATTGAG GGAAAGCATG TCTGCTGGGT GTGACCATGT 1260  
 TTCTCTCAA TAAAGTCCC CTGTGACACT C

Seq ID NO: 109 Protein sequence:  
 Protein Accession #: NP\_006133.1

1 11 21 31 41 51  
 MERASLIQKA KLAQAERYE DMAAFMKGAV EKGEELSCEE RNLLSVAYKN VVGQRAAWR 60  
 VLSSIEQKSN BEGSEKGPV VREYREKVT ELQGVCDTVL GLLDHLIKE AGDAESRVFY 120  
 LKMGDYRY LAEVATGDDK KRIIDSARSA YQRAMDISK EMPPTNPIRL GLALNFSVFH 180  
 YEIANSEEA ISLAKTFDE AMADLHTLSE DSYKSTLIM QLLRDLNLTLM TADNAGEEGG 240  
 EAPQSPQS

Seq ID NO: 110 DNA sequence  
 Nucleic Acid Accession #: NM\_000695  
 Coding sequence: 407-1564

1 11 21 31 41 51  
 CACGAGTGG TTTGGGAGCT GCCAGTCTCC TGGGAGGATC GCAGTCAGCA GAGCAGGGCT 60  
 GAGGCCTGGG GGTAGGAGCA GAGCCTGCGC ATCTGGAGGC AGCATGTCCA AGAAAGGGAG 120  
 TGGAGGTGCA GCGAAGGACC CAGGGCAGCA GCCCACGCTG GGGATGGACC CCTTCGAGGA 180  
 CACACTCGCG CGGCTGCGTG AGGCCTTCAA CTGAGGCGCG AOCGCGCGCG CCGAGTTCOG 240  
 GGCTGCGCAG CTCAGGGGCC TGGGCCACTT CCTTCAAGAA AACAGCAGC TCTGCGCGA 300

Seq ID NO: 111 Protein sequence:  
Protein Accession #: NP\_000686

55 Seq ID NO: 112 DNA sequence  
Nucleic Acid Accession #: NM\_004456  
Coding sequence: 58-2298

231

CCCTGTGATC ATCCACGCA GCCTGTGAC AGTTCGTGCC CTGTGTGAT AGCACAAAT 1680  
 TTTTGTGAAA AGTTTGTGCA ATGTAGTCA GAGTGTCAAA ACCGCTTTCC GGGATGCCGC 1740  
 TGCAAGCAC AGTGCAACAC CAAGCAGTGC CGTGCTACC TGGCTGTCCG AGAGTGTGAC 1800  
 CCTGACCTCT GTCTTACTTG TGGAGCGCT GACCATTTGG ACAGTAAAAA TGTGTCTGTC 1860  
 AAGAACTGCA GTATTACGG GGGCTCCAAA AAGCATCTAT TGCTGGCACC ATCTGACGTG 1920  
 CGAGGCTGGG GGATTTTAT CAAAGATCCT GTGCAGAAAA ATGAATTAT CTGAGAAATAC 1980  
 TGTGGAGAGA TTAATTTCTA AGATGAAGCT GACAGAAGAG GGAAGTGTA TGATAAATAC 2040  
 ATGTGCAGCT TTCTGTTCAA CTGAACAAT GATTTGTGG TGGATGCAAC CCGCAAGGGT 2100  
 AACAAAAATC GTTTTGCAAA TCATTCCGTA AATCCAACT GCTATGCAAA AGTTATGATG 2160  
 GTTAACGGTG ATCAGAGAT AGGTATTTTT GCCAAGAGAG CCATCCAGAC TGGCGAAGAG 2220  
 CTGTTGTGTT ATTACAGATA CAGCCAGGCT GATGCCCTGA AGTATGTCGG CATCGAAAGA 2280  
 GAAATGAAA TCCCTTGACA TCTGTACCT CCTCCCTCTC CTCTGAAACA GCTGCCTTAG 2340  
 CTTCAGGAAC CTCGAGTACT GTGGCAATT TAGAAAAAGA ACATGCAGTT TGAATTTCTG 2400  
 AATTGCAAA GTACTGTAG AATAATTTAT AGTAATGAGT TTAAAAATCA ACTTTTATT 2460  
 GCCTTCTCAC CAGCTGCAAA GTGTTTGTGA CCAGTGAATT TTTGCAATAA TGCAATATGG 2520  
 TACATTTTTC AACTTTGAAT AAAGAATACT TGAACCTGAA AAAAAAATA AAAAAA

Seq ID NO: 113 Protein sequence:  
 Protein Accession #: NP\_004447

1 11 21 31 41 51  
 MGQTGKSEK GPVCRKRKVK SEYMRLRLQK RFRRADDEVKS MFSSNRQKIL ERTEILNQEW 60  
 KQRRIQPVHI LTVSVSLRGT RECVTSDDL DFPQVPLKLT LNAVASVPIM YWSPLQGNF 120  
 MVEDETVLWN IPYMGDEVLD QDGTPIELI KNYDGKVGHD RECGFINDEI FVELVNALQG 180  
 YNDDDDDDG DDPEREREKQ KLEDEHRDDK ESRPPRKFPF DKILEATISSM FPDGTAEBEL 240  
 KEKYKELTEQ QLPALPPEC TPNIDGNPAK SVQREQSLES FHTLFCRRCP KYDCFLHPFH 300  
 ATPNTYKRNK TETALDNKPC GPQCYQHLEG AKESFAAALTA ERIKTPPKRP GRRRRGRLLN 360  
 NSSRPSTPTI NVLESKDTDS DREAGTETGG ENMDKEEEK KDETSSESEA NSRCQTPIM 420  
 KPNIEPPENV EWSGAEASMP RVLIGTYIDN FCAIARLIGT KTCRQVVEFR VKESSIIAPA 480  
 PAEDVDTPPR KKKRKRERLWA AHCRIQLK DGSSNHVINY QPCDHPRQPC DSSCPVIAQ 540  
 NPKCFKQCS SEQNRPFGC RCKAQNTKQ CPCYLAVREC DPDLCLTCGA ADHWDNRNVS 600  
 CKNCISIQRS KHELLLAPSD VAGWGFIDK PVQKNEFISE YCGEIIISQDE ADRRQKVYDK 660  
 YMCSPFLNIN NDFVVDATRK GNKIRFANHS VNPENYAKVM MVNGDHRIGI FAKRAIQTGE 720  
 ELFDYRYSQ ADALYVGIE REMEIP

Seq ID NO: 114 DNA sequence  
 Nucleic Acid Accession #: NM\_001827  
 Coding sequence: 96-335

1 11 21 31 41 51  
 AGTCTCCGGC GAGTTGTTGC CTGGGCTGGA CGTGGTTTGG TCTGCTGCGC CGCTCTTCG 60  
 CGCTCTCGTT TCATTTCTCG CAGCGCGCCA CGAGGATGGC CCACAAGCAG ATCTACTACT 120  
 CGGACAACTA CTTGACGAAA CACTACGAGT ACCGGCATGT TATGTTACCC AGAGAACTTT 180  
 CCAACCAAGT ACCTAAACT CATCTGATGT CTGAAGAGGA GTGGAGGAGA CTGGGTGTC 240  
 AACAGAGTCT AGGCTGGGTT CATTACATGA TTCATGAGCC AGAACCCAT ATTCTTCTCT 300  
 TTAGACGACC TCTTCCAAA GATCAACAAA AATGAAGTTT ATCTGGGAT CGTCAAACTCT 360  
 TTTTCAAAAT TAATGTATAT GTGTATATAA GGTAGTATTC AGTGAATACT TGAGAAATGT 420  
 ACAAACTCTT CATCCATACC TGTGCATGAG CTGTATTCTT CACAGCAACA GAGCTCAGTT 480  
 AAATGCAACT GCAAGTAGGT TACTGTAAGA TGTTTAAGAT AAAAGTTCTT CCAGTCAGTT 540  
 TTTCTCTTAA GTGCTGTTT GAGTTTACTG AAACAGTTTA CTTTGTGTTCA ATAAAGTTTG 600  
 TATGTTGCAT TAAAAAATA AAAAAA

Seq ID NO: 115 Protein sequence:  
 Protein Accession #: NP\_001818

1 11 21 31 41 51  
 MAHKQIYYSD KYFDEHYEYR HVMLPRELSK QVPKTHLMSE EEWRLIGVQQ SLGWVHYMIH 60  
 EPEPHILLFR RPLPKDQK

Seq ID NO: 116 DNA sequence  
 Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
 TCAGACCTCA TGAGTCACTT GGAATCTTGA GCCACCTCTG GGGGTGGAGT CTCTCTCCTG 60  
 GCATCTGGAC CTTGGTGTCT ATCGACGAAG CTTGGGTGGG GCTCTTAGCT GCTATGTGCA 120  
 AGAGGTGTGT TCCAGGAAA GCCCTATCT CTCTGCAGAG GTCAAGTGAA AGCGACGGCC 180  
 GCAGCCAACA GAGTTCAAAA TGCAGGCTTG GAAAGTACAG GGGGCTCTGT GGAGGATGGG 240  
 AAGGACTGAT CCACATTCCC ACCAGGAAGT TTAGCAGAAC CCCGCGTGC CAATGGGACC 300  
 CCTTGGAGAG ACCTGGCTCA GGCTGGACCA CCTCTTGAGA GGGAGGAGCT CTGGATTGTA 360  
 TCAAGAAATC TTGCTGAGC ATGGTGCCCTC ATGCCTATAA TACCAACACT TTGGGAGGCC 420  
 AGTGTGGGAG GATCTCTTGA GCCAGGAGT TCAAGACTAG CTGGGGCAAC ACAGAGAGAA 480  
 CCCATCTCTA AAATAATAAT AATAATAAAA TAAAAAATA GCAGGGCATG GTGGCATGTG 540  
 CCTGTAGTTC CAGCTACCCA GGAGGCTGAG GCAAGAGGAT GGCTGGAGCC TGGGATGTTG 600  
 AGGCTGCAAT GAAGTGTGAT TACCCCACTG CACTCCAGCC TGGGCAAAAG AGCGAGAGAA 660  
 CTTGTCTCAA ATAAATAAAT TAATAATAAT CTATTTTTGG AGAATAAAGA GACCTCTGGA 720  
 TTTGAGGTGC CATTTGGGTA GAAAGAAAAG ACGTTTACAC CGAGAAATAG TCTGTGTTGC 780  
 CCTGAAGGAG CAGAGGGATG CATCGCTGGA GGTGACCTAC AGTTGAAGAA GACTCATTAT 840  
 GACAGACCTT GTCTTCTTC CTGTGGAAA GTGTTCTCTC TGCTGCTACT GCTCATGAGA 900  
 CTCTTCCCCC TCCTGTGCC AGGGAACCAA AGGGCTTTCT ACCACACCTT TTCTTGGCCC 960  
 CGGCTCTCCA TGTGCTGCT GCCTTTGTAC TCAGCAATTC TTGTTGTCTC CATTATCTTC 1020  
 CAGCGGGATA CAGAGTGAAT AGTTAACCACT ACTTAGGTCA AATAGGATCT AAATTTTGT 1080  
 TCCTGCTCCG TGTAAGAGG CAGTGTGTTG TGTGTTGCAA GCAGCCTTGG AATAGTAAT 1140



## WO 02/086443

PCT/US02/12476

CTTCTCATTT GTTGGGATC TGGCCACCAA GTTCCAGAAT GATACACGGA TCAGTGCAGA 1200  
AGTTTCATCAG GCTCTCGGAC CTTAGGGCTG TTGGAGAAGG CTTTCAGCAGC AGAACTGATG 1260  
GTGAAGGCTC GTGTTCTCCA TCCTCAACTT TCTTTGCTTC GATCATACAC AAGAATACAT 1320  
TTGGAAGGGC AAAAATAGAA CACTGTGTTT CATTGCAGCC GTGTTTGTG ACACAGATGC 1380  
ACAGTCTGCT GTGAAGACCT TCTCTCAAGT GGCATTGGGG AGTCCATGCC AGATCATGGT 1440  
GCTTCATGAG AGACTGACAC CTATCAGGGG TTGTGGCACT TAGTGAGGAC TCTCTCTCCC 1500  
CAGTGTGTGC TGATGACACA TACACACCTG ACAATAGCTT GAGTCTTCTC TGTTCCTTTT 1560  
ACTCTGTAGC CAACATACAC ATGATTTAAA ACCCTTTCTA AATATCTATC ATGGTTCATC 1620  
CTGTGCCAAA TGCAGAGTCA GAGCTATTG TACTTCATTA TTATTTCCAA GCGAATAGT 1680  
TGGCTTTCTT TTTGCAAAAA TAATTAAAGT TTTTGTATGT TGCAAAAAAA AAAAAAATAA 1740  
AAAAAATAAA

Seq ID NO: 117 DNA sequence  
Nucleic Acid Accession #: BC012178.1  
Coding sequence: 204-2285

1 11 21 31 41 51  
CTTCTCTCCC GCGCGCTGG GCGCGGCGCT CCGCTGCTGT TGCTCCATTC GCGCTTTTTC 60  
TGGGCGCTGG CTCCTCTCCG CTGCGGCTG CTCTCGGACC AGGCCTCCTT CTCAACCTCA 120  
GCGCGCGGGG CCGACCCCTC CCGCACCCCT CCGCCCCGTC TCGTACTGTC GCGCTCACCG 180  
CCGCGGCTCC GGCCTCGGCC CCGATGGCTC TGTGCAACGG AGACTCCAAG CTGGAGAAATG 240  
CTGGAGGAGA CTTTAAGGAT GGCACCAACC ACTATGAAGG AGCTGTTGTC ATTCTGGATG 300  
CTGGTGCTCA GTACGGGAAA GTCATAGACC GAAGAGTGAG GGAAGTGTTC GTGCAGTCTG 360  
AAATTTTCCC CTGGAAACA CCAGCATTG CTATAAGGA ACAAGGATTC CGTGCTATTA 420  
TCATCTCTGG AGGACCTAAT TCTGTGTATG CTGAAGATGC TCCTGCTTTT GATCCAGCAA 480  
TATTCATAT TGSCAAGCCT GTTCTTGAAA TTGTCTATGG TATGCAGATG ATGAATAAGG 540  
TATTGGAGG TACTGTGCAC AAAAAAAGT TCAGAGAAGA TGGAGTTTTC AACATTAGTG 600  
TGGATAATAC ATGTTTCTTA TTCAGGGGCC TTCAGAAGGA AGAAGTGTGT TTGCTTACAC 660  
ATGGAGATAG TGTAGACAAA GTAGCTGATG GATTCAAGGT TGTGGCACGT TCTGGAAAAA 720  
TAGTAGCAGG CATAGCAAAAT GAATCTAAAA AGTTATATGG AGCACAGTTC CACCCTGAAG 780  
TTGGCCTTAC AGAAAAATGA AAGATAATAC TGAAGAATT CTCTTATGAT ATAGCTGGAT 840  
GCAGTGAAGC CTTACCGCTG CAGAACAGAG AACTTGAGTG TATTGAGAG ATCAAAGAGA 900  
GAGTAGGCAC GTCAAAAGTT TTGTTTTC TCAGTGGTGG AGTAGACTCA ACAGTTTGTA 960  
CAGCTTTGCT AAATCGTGTCT TTGAACCAAG AACAGTCTAT TGCTGTGCAC ATTGATAATG 1020  
GCTTTATGAG AAAACGAGAA AGCCAGTCTG TTGAAGAGGC CCTCAAAAAG CTGGGAATTC 1080  
AGGTCAAAGT GATAATGTCT GCTCATTCTT TCTACATAGG AACACAACCT CTACCAATAT 1140  
CAGATGAAGA TAGAACCCCA CGSAAAAGAA TTAGCAAAAC GTTAAATATG ACCACAAGTC 1200  
CTGAAGAGAA AAGAAAAATC ATTGGGGATA CTTTGTGTTA GATTGCCAAT GAAGTAATG 1260  
GAGAAATGAA CTTGAAACCA GAGGAGGTTT TCCTTGCCCA AGTACTTTA CGGCTGATC 1320  
TAATTGAAAG TGATCCCTT GTTGCAAGTG GCAAAGCTGA ACTCATCAA ACCCATCA 1380  
ATGACACAGA GCTCATCAGA AAGTTGAGAG AGGAGGGAAA AGTAATAGAA CCTCTGAAG 1440  
ATTTTCATAA AGATGAAGT AGAATTTTGG GCAGAGAACT TGGACTTCCA GAAGAGTTAG 1500  
TTTCCAGAGA TCCATTTC CAAGCTGCTG TGGCAATCAG AGTAATATGT GCTGAAGAAC 1560  
CTTATATTGG TAAGGACTTT CCTGAAACCA ACAATATTTT GAAATAGTA GCTGATTTT 1620  
CTGCAAGTGT TAAAAAGCCA CATACCTAT TACAGAGAGT CAAAGCCTGC ACACAGAG 1680  
AGGATCAGGA GAAGCTGATG CAAATTACCA GTCTGCATTC ACTGAATGCC TTCTGCTGCTG 1740  
CAATTAAAAA GTAGGTGTGT CAGGCTGACT GTGCTTCTTA CAGTTACGTC TGTGGAATCT 1800  
CCAGTAAAGA TGAACCTGAC TGGGAATCAC TTATTTTCTT GGCTAGGCTT ATACCTCGCA 1860  
TGTGTACAAA CGTTAACAGA GTTGTTTATA TATTGGCCCC ACCAGTTAAA GAACCTCCTA 1920  
CAGATGTAC TCCACTTTC TTGACACAG GGGTGTCTAG TACTTTACGC CAAGCTGATT 1980  
TTGAGGCCCA TAACATTCTC AGGGAGTCTG GGTATGCTGG GAAATCAGC CAGATGCCCG 2040  
TGATTTTGAC ACCATTACAT TTTGATCGGG ACCCACTTCA AAAGCAGCCT TCATGCCAGA 2100  
GATCTGTGTT TATTCGAACC TTTATTACTA GTGACTTCAT GACTGTGATA CCTGCAACAC 2160  
CTGGCAATGA GATCCCTGTA GAGGTGTTAT TAAAGATGTT CACTGAGATT AAGAAGATTC 2220  
CTGGTATTTT TCGAATTATG TATGACTTAA CATCAAAGCC CCGAGGAAGT ACTGAGTGGG 2280  
AGTAATAAAC TTCTGTTCT ATTAATAA

Seq ID NO: 118 Protein sequence:  
Protein Accession #: AAH12178.1

1 11 21 31 41 51  
MALCNGDSKL ENAGGLDKG HHHYEGAVVI LDAGAQYQKV IDRRVRELFV QSEIFPLETP 60  
APAIKEQGR AIIISGGPNS VYADAPWFD PAIFTIGKPV LGICYQMOM NKVFGGTVHK 120  
KSVREDFVN ISVDNTCSLF RGLQKESVVL LTHGDSVDKV ADGFKVARS GNIVAGIANE 180  
SKKLYGAQFH PEVGLTENGG VILKNFLYDI AGCSGTFTVQ NRELECIREI KERVGTSKVL 240  
VLLSGGVDS VCTALLNRAL NQEQVIAVHI DNGFMRKRES QSVEEALKKL GIQVKVINAA 300  
HSFYNGTTL PISDEDRTPR KRISKTLNMT TSPREKRKII GDTFVKIANE VIGEMNLKPE 360  
EVFLAQGLTR PDLIESASLV ASGRAELIKT HHNDTELIRK LREEGKVIPE LKDFHKDEVR 420  
ILGRELGLPE ELVSRHPPFG PGLAIRVICA EEPYICKDFP ETNNILKIVA DFSASVKKPH 480  
TLQORVACT TEEDQEKLMQ ITSLSLNAF LLPIKTVGVG GDCRSYSYVC GISSKDEPDW 540  
ESLIFLARI PRMCNVRNV VYIFGPFVKE PPTDVTPTFL TTGVLSTLRQ ADPEAHNLR 600  
ESGYAGKISQ MPVILTPLEF DRDPLQKQPS QRSVVIRT ITSDFMGTGIP ATPGNEIPE 660  
VVLKMTETIK KIPGSRIMY DLTSKPPGTT EWE

Seq ID NO: 119 DNA sequence  
Nucleic Acid Accession #: NM\_006500.1  
Coding sequence: 27..1967

1 11 21 31 41 51  
ACTTGGCTCT CGCCCTCCGG CCAAGCATGG GGCTTCCAG GCTGGTCTGC GCCTTCTTGC 60  
TGCCGCGCTG CTGCTGCTGT CCTCGGCTCG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120  
CGCCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCTCTCTCC 180  
AGTCCCAAGG CAACCTCAGC CATGTGACT GGTTTTCTGT CCACAGGAG AAGCGGACGC 240



TCATCTCCG TGTGGCGCC GGCACGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC 300  
TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCTGACTCA AGTCAACCCC CAAGACGAGC 360  
GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGGC 420  
TCTACAAAGC TCCGGAGGAG CCAAAACATCC AGGTCAACCC CTTGGGCATC CCTGTGAACA 480  
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGGTACCCC ATTCTCAAG 540  
TCATCTGGTA CAGGAATGGC CGGCTCTGA AGGAGGAGAA GAACCGGGTC CACATTCAGT 600  
CGTCCCAGAC TGTGGAGTCG AGTGGTTTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660  
TGGTTAAAGA AGACAAAGAT GCCCAGTTTT ACTGTGAGCT CAACCTACCG CTGCCAGTG 720  
GGAAACCAT CAT GAAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCG ACAGAAAAG 780  
TGTGGCTGGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCGTG GAAATCAGGT 840  
GTTTGGCTGA TGGCAACCTT CCACCACTT TCAAGCATCAG CAAGCAGAAC CCCAGCACCA 900  
GGGAGGCAGA GGAAGAGACA ACCAACGACA ACGGGTCTCT GGTGCTGAG CCTGCCCGGA 960  
AGGAACACAG TGGCGCTTAT GAATGTGAGG CTGGAACCTT GGACCATG ATATCGCTGC 1020  
TGAGTGAACC ACAGGAACCTA CTGGTGAAC ATGTGTCTGA CGTCCGAGTG AGTCCCGCAG 1080  
CCCTCGAGAG ACAGGAAGGC AGCAGCCTCA CCTGACCTG TGAGGCAGAG AGTAGCCAGG 1140  
ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAGGT GCTGGAAGG GGGCCTGTGC 1200  
TTCAGTTGCA TGACCTGAAA CGGAGGCGAG GAGGCGGCTA TCGTGTGCTG GCGTCTGTGC 1260  
CCAGCATACC CGGCTGTAAC CGCACACAGC TGGTCAAGCT GGCCATTTTT GCGCCCTCTT 1320  
GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAGAGAA TATGTTGTG AATCTGTCTT 1380  
GTGAAGCGTC AGGGCAGCCC CGGCCACCA TCTCTGTGAA CGTCAACGGC ACGGCAAGTG 1440  
AACAAAGACA AGATCCACAG CGAGTCTGGA GCACCTGAA TGTCTCTG AGCCCGGAGC 1500  
TGTGGAGAC AGGTGTGAAA TGCAAGGCTT CCAACGACCT GGGCAAAAAC ACCAGCATCC 1560  
TCTTCTCGGA TCTGGTCAAT TTAACCAACC TCACACAGA CTCACACACA ACCATGGCC 1620  
TCAGCACTTC CACTGCCAGT CCTCATACCA GAGCCAACAG CACCTCCACA GAGAGAAAGC 1680  
TGCCTGGAGC GGAGAGCGGG GGGTGGTCA TCGTGGCTGT GATTGTGTGC ATCTGGTTC 1740  
TGGCGGTGCT TGGCGCTGTC CTCTATTTCC TCTATAAGAA GGGCAAGCTG CCGTGCAGGC 1800  
GCTCAGGAA GCAGGAGATC ACGCTGCCCC CGTCTCGTAA GACCGAATCT GTAGTTGAAG 1860  
TTAAGTCAGA TAAGTCCCA GAAGAGATGG GCCTCTGCA GGGCAGCAGC GGTGACAGA 1920  
GGGCTCCGG AGACACAGGA GAGAAATACA TCGATCTGAG GCATTAGCCC CGAATCACTT 1980  
CAGCTCCCTT CCTGCTGAG ACCATTCCCA GCTCCCTGCT CACTCTTCTC TCAGCCAAAG 2040  
CCTCCAAAGG GACTAGAGAG AAGCCTCTCT CTCCTCTCAC CTGCACACCC CTTTCAGAG 2100  
GGCCTAGGG TTAGGAGCTG AGGACCTCAC TTGGCCCTGC AAGCGCTTT TCAGGAGACA 2160  
GTCCACCAAC ATCTCTCCA CGTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCCAGTCTC 2220  
CCGAGCGGT AGGAGAGTTT CTTCAGAAC GTTTTTTTC TTACACACA TTATGGCTGT 2280  
AAATACCTGG CTCTGCGCAG CAGCTGAGCT GGGTAGCTCT TCTGAGCTGG TTTCTGCCC 2340  
CAAAAGCTGG CTTCACCAT CCAGGTGCAC CACTGAAGTG AGGACACACC GGAGCCAGGC 2400  
GCCTGCTCAT GTTGAAGTGC GCTGTTCACA CCGCTCCGG AGAGCACCCC AGCGGCATCC 2460  
AGAAGCAGCT CAGGTGTGTC TGCCACCAAC CTCTGCTCG CCTTTTCAA GTCTCTGTG 2520  
ACATTTTTTC TTTGGTCAGA AGCCAGGAAC TGGTGTCTT CTTTAAAGA TACGTGCCGG 2580  
GGCCAGGTGT GGTGGCTCAC GCGTGAATC CCAGCACTTT GGGAGCCGA GCGGGGCGGA 2640  
TCACAAAGTC AGGAGCAGAC CATCTGGCTT AACACGCTGA AACCTGTCT CTAATAAAAA 2700  
TACAAAAAAA AATTAGCTAG GGTGAGTGT TGGACCTAT AGTCCAGCT ACTCCGAAGG 2760  
CTGAAGCAGG AGAATGTTAT GAATCCAGGA GGTGGAGCTT GCAGTGAGCC GAGACCGTGC 2820  
CACTGCATCT CAGCCTGGGC AACACAGCGA GACTCCGTCT OGAGGAAAAA AAAAGAAAA 2880  
ACGCGTACCT CGCGTGAGA AGCTGGGCGC TGTTTTCGAG TTCAAGTGAA TTAGCTCAA 2940  
TCCCGGTGTT CACTGTGCTC CATAGCCCTC TTGATGATC ACGTAAACT GAAAGGCAGC 3000  
GGGAGCAGA ACAAAGTAG GTCTACACTG TCCTTCATGG GGATTAAGC TATGTTTATA 3060  
TTAGACCAA ACTTCTCAA ACCAAGCTCA GGGCCCAAC CCTAGAAGG CCCAATGAG 3120  
AGAATGGTAT TTAGGAGTGG AAAACGGGGC CTGGCTAGAG CTTCCGGGTG GTGTGTCTGT 3180  
CTGTGTGTAT GCATACATAT GTGTGTATAT ATGGTTTTGT CAGGTGTGTA AATTGCAAA 3240  
TTGTTTCTT TATATATGTA TGTATATATA TATATGAAA TATATATATA TATGAAAAAT 3300  
AAAGCTTAAT TGTCCCAGAA AATCATACAT TGCTTTTTTA TTCTACATGG GTACCAACAG 3360  
AACCTGGGGG CCTGTGAAC TACAACCAA AGGCACACAA AACCGTTTCC AGTTGGCAGC 3420  
AGAGATCAGG GGTACTCTCT GCTTCTGAGC AAATGGCTCA AGCTCTACCA GAGCAGACAG 3480  
TACCTCTACT TTTACAGCAG AAAAGTCCC GTATGACGCA GCACGAAGG CCTGGCAGGC 3540  
TGTTAGCAGG AGCTATGTCC CTTCCTATCG TTTCCGTCCA CTT

Seq ID NO: 120 Protein sequence:  
Protein Accession #: NP\_006491.1

1 11 21 31 41 51  
MGLPRLVCAF LLAACCCCPR VAGVPGEAEQ PAPELVEVEV GSTALLKQGL SQSQGNLSHV 60  
DWFSVHKEKR TLIFRVROGG QOSEPGEYEQ RLSLQDRGAT LALTQVTPQD ERIFLCQGRK 120  
PRSQEYRIQL RVYKAPBEPN IQVNPLGIPV NSKEPEEVAT CVGRNGYPIP QVIWYKNGRP 180  
LKEEKNRVHI QSSQTVESSG LYTLQSLKLA QLVKEDKDAQ FYCELYNRLP SGNEHMKESRE 240  
VTVPVFYPTE KVNLEVPVG MLKBDGDRVZI RCLADGNPPP HFSISKQNPFS TREAEZETIN 300  
DNGVLVLBPA RKEHSGRYEC QAWNLDTMIS LLSEPQELLV NYVSDVRVSP AAPERQEGSS 360  
LTLTCEAESS QDLFQWLRLE ETDQVLERGP VLQLHDLKRE AGGGYRCVAS VPSIPGLNRT 420  
QLVKLAIFGP PMAFRERKV WVKENMVLNL SCEASGHPRP TISWNVNGTA SEQDDPQORV 480  
LSTLNLVLT ELLETGVECT ASNDLGKNTS ILPLELVNLT TLTPDSNTTT GLSTSTASPH 540  
TRANSTSTER KLPEPESRGV VIVAVIVCIL VLAVLGAVLY FLYKKGKLPC RRSKQBITL 600  
PPSRKTELTV EVKSDLPPEE MGLLQSSGD KRAPGDQGER YIDLRR

Seq ID NO: 121 DNA sequence  
Nucleic Acid Accession #: NM\_018306  
Coding sequence: 60-671

1 11 21 31 41 51  
ATAGTCTACA CAGAGCTCCC CTTGCTGCCC AGACAAGCTG AAGGACCACA GGAAAAGCCA 60  
TGGAGACTTC AGCATCTCTC TCCAGCCTC AGGACAACAG TCAAGTCCAC AGAGAAACAG 120  
AAGATGTAGA CTATGGAGAG ACAGATTTC ACAGACAAGA CCGGAAGGCT GGACTCTTTT 180  
CCCAAGAAC ATATGAGAGA AACAAGTCTT CTTCCTCTC CTCTCTTCC TCCTCATCTT 240  
CCTCATCTTC TTTATCTCTC TCCTCTCTCAG GTCTGCGGCA TGGGAGCCT GACGTTTTGA 300

5  
10  
15  
20

```

AGGATGAGCT TCAACTCTAT GGAGATGCTC CTGGAGAGGT GGTACCCCTCT GGGGAATCAG 360
GACTCCGAGG GAGAGGCTCT GACCCAGCAA GTGGAGAGGT GGAGGCTCTCT CAGTTAAGAA 420
GACTGAATAT AAAGAAAGAT GATGAGTTTT TCCATTTCTG CTCTCTGTGC TTTGCCATCG 480
GGGCGCTTCT GGTGTGTAT CACTATTACG CAGACTGGTT CATGTCTCTT GGGGTGCGCC 540
TGCTCACCTT CGCCTCCCTG GAAACCGTTG GCATCTACTT CGGACTAGTG TACCGTATCC 600
ACAGGCTCCT CCAAGGCTTC ATCCCTCTCT TCCAGAAGTT TAGGCTGACA GGGTTCAGGA 660
AGACTGACTG AGGCCATCTC CAGGTGGGCA GCAGAGGCAG GCCCCAGTGT GACCACCACT 720
GCGACCCCTG AGCCCCAAGG GGCAGAGCAG CATTCTGAGA GACGCACAGG AGACCAAGCC 780
AGACCAATAA ACAGAACACT TTTCTTCCA TGTGCTCTGA ATGTTGGCAC CAGCCCGGGC 840
AGGGGCATCT CATTTGGGCA GTACTGCTGT GCAACCCAGC TGCAAGGATG GAAGGCAGAG 900
GGTGGGTGTG GGGCCTGAGG CTTACAGTA CCTGGACCAG CAGGAAGATT CTGGGAGGTC 960
ATGCTCTCA GAGGACAGCA AGGGACCTTG AGCTCTGCAA GCTGTGATCT GTCTGGGTC 1020
ATGTTTTTTC TCAATCCCA GGCTATCTGC ATGCGCTCTC AGGTGCTACC GAGCCATCCT 1080
GGGAGAGATG GATGGTCCAC TGCTTTGAGG CAGGGAGCCA TCGGGCTGGG GCCCTTTGGT 1140
GAACCTGATG CAGGTAAGAT GCTGAGGACT AAAACCATTT TTTTTCACCC CAAAAAATAA 1200
GCAGGAAATA TGCATCATCAG AAATAAATG GCAGCCAGGC ATGGGGGCTC ACGACTGTAA 1260
TCTCGCACT TTGGAGGCTC CAGGCTAAGG GTCCGCTGAA GCTGAGAGTT CAAGACCAAC 1320
CTGGGCAACA TAGTGAGACC CCAATCTCTA CAATTTTTTT TTAATGACCA AATGTGGCGG 1380
TACATACCTG TACATACCTG CGGTTCCAGC TACTCAAGAG GCTGAGGCAG GAGGACTGCT 1440
TGAGCCAGG AGTTCAGGC TGCACTGAGG TACGATCAAG CCACTGCACT CCAGCCTGGG 1500
CGACAGAGCA AGATCGTTTC TCTAAAAAT

```

Seq ID NO: 122 Protein sequence:  
Protein Accession #: NP\_060776

30

```

1 11 21 31 41 51
| | | | |
METSASSSQP QNDSQVHRET EVDVYGETDF HKQDGKAGLP SQEQYERNKS SSSSFSSSSS 60
SSSSSSSSSS GPGHGEPDVL KDELQLYGDA PGEVVPSCGS GLRRRGSDPA SGEVEASQLR 120
RLNIKKDDEF FHFVLLCPAI GALLVCYHYI ADWFMSLGVG LITFASLETV GYFGLVYRI 180
HSV LQGFIPL FQKFR LITGFR KTD

```

Seq ID NO: 123 DNA sequence  
Nucleic Acid Accession #: BC022542  
Coding sequence: 243..896

40  
45  
50  
55  
60  
65  
70

```

1 11 21 31 41 51
| | | | |
ACTTGGTCCC AGCCGATAAA TCTGGGGCAG CGCGCGGTAG GAGCTGCGGG CGGCCAGGCC 60
CCTTCTCTCG TCCGCACTTG GCGCCGCGCG CCGCTCTCGG GCGTCCGGCT TCCGCGTCC 120
TGCGGGCTCG GGTGGCGCGG GTTCGGGGCG CGCGCTGCTG GCTCTCTCGG GCGGCGACGG 180
GGCTCACGCG GCGGCGCGCC ACGGCCCTCA CCGCGCGCGG CTCTGACGCC GGCATAAGGG 240
CCATGTGTTT TGAATATTAT TTGAGGCAAG AAGTTTGTAA AGATGGTTTC CACAGAGACC 300
TTTTAATCAA AGTGAAGTTT GGGGAAAGCA TTGAGGACTT GCACAGTGC CGTCTCTTAA 360
TTAAACAGGA CATTCTCTCA GGACTTTATG TGGATCCGTA TGAGTTGGCT TCATTACGAG 420
AGAGAAACAT AACAGAGGCA GTGATGGTTT CAGAAAAATT TGATATAGAG GCCCCTAACT 480
ATTGTGTCAG GAGACTCTGA GTTCTCATT ATGCCAGACG AGATTACAG TGCAATTGACT 540
GTTTTCAAGC CTTTTTGCTT GTGCACTGCC GCTATCATCG GCGGCACAGT GAAGATGGAG 600
AAGCCTCAGT TGTGTTCAAT AACCCAGATT TGTGTATGTT TTGTGACCAA GAGTTCCTGA 660
TTTTGAATG CTGGGCTCAC TCAGAAAGTG CAGCCCTCTG TGCTTTGGAT AATGAGGATA 720
TATGCCAATG GAACAGATG AAGTATAAAT CAGTATATAA GAATGTGATT CTACAAGTTC 780
CAGTGGGACT GACTGTATAC ACCTCTCTAG TATGTTCTGT GACTCTGCTC ATTACAATCC 840
TGTGCTCTAC ATGTACTCTT GTAGCAGTTT TCAAAATATG CCAATTTTCC CTATAAGTTT 900
TATGTAGTTA ATGTCTCTCT AGAAACCTAA ATAAGATCTA TTAATTTCTG ACGAGAGGTG 960
TTCTTCTAGA ATTAATTAAT TTTATCTTTT GTCTTCATT GTGGCCAAA TTAGTTTAC 1020
TAGAGAAAT TTGGGATCAT TCTCAGCTAA TTCCAAATG TAGTCTCTA TTGCAATGAT 1080
CCTTGGTAAT CCTCAAGCAT CAGATGCCAT AAGGGGAAAC TTAATTTCTG TAAATTAATG 1140
TTTATTTTGT GAGAAGTGAC TTTATCTTCA TTTGGGGTAG AAAAATTATT TCTTTATGTA 1200
GTAGAGACAA ATATTCTTCA TTTTGCAAGT ACTTTCAATT TAAGCTACAA ATTGAGAAAA 1260
CGTTATATAA TAAGATATAA ATAGGCCAGG CACAGTGGCT CACACCTGTA ATCCAGCAC 1320
TTTGGGAGGC CGAGGTGGGC GGATCACCAG AGGTCAAGAG TTTGAGACCA GCTTGGTGAA 1380
ACCCTGTCTC TACTAAAAAT ACAAAGTTA GCTGGGGCTG GTGGTGGGCA TCTGTAGTCC 1440
CAGCTAATTG GAAGGGTGAG GCGGGAGGAT CGCTTGAACC TGGGAGGCGG AGGTTCCAGA 1500
GAGCCCAAGT CGCACCCTG CACTACAGCC TGGGCGACAG AACGAGACCC TGTCTCCAAA 1560
GGAAAAACAA AAAAGAGAA TAAATAAATT TGGATGAAAA TCATGTTTAT TTAATAGTA 1620
ATGTCTAGAG ACTATTAAAG ATGTGCCAGA GTTCAATGA AAATCATTAA AGTAGGACAG 1680
CTAAGAAATT AATATTAATA TAAAAATTAT TGATAATCTT AAATTATTGA TTAATCCTTA 1740
ACGCACTCCA TTCTCTTTT ACATTTTATC ATGTTTCTTT TGAATATATG AATTGGCAAA 1800
GGACTTGATG AAATCTAGTA CTAAGATTG GTACAGAGTA TGTGAGGAAG ACAACTCAGA 1860
TTGCCATTTT AAATAAAGTT GTACATGAAC AAAAAAAAAA AAAAAA

```

Seq ID NO: 124 Protein sequence:  
Protein Accession #: AAH22542

80

```

1 11 21 31 41 51
| | | | |
MCSEILRQE VLKDGPHRDL LKVKFGESI EDLHLCRLLI KQDIPAGLYV DPYELASLRE 60
RNITEAVMVS ENFDIEAPNY LSKSEVLIY ARRDSQCIDC FOAFLPVHCR YHRPHSEDEGE 120
ASIVVNNPDL LMFCDQAGSR RMIRFRFDSF DKTIEFPILK CWAHSEVAAP CALENEDICQ 180
WNMKYKSVY RNVILQVPVG LTVHTSLVCS VTLLITLILCS KKKKK

```

Seq ID NO: 125 DNA sequence  
Nucleic Acid Accession #: NM\_004994.1  
Coding sequence: 20..2143

1 11 21 31 41 51  
 AGACACCTCT GCCCTACCCA TGAGCCTCTG GCAGCCCTG GTCTGTGTGC TCCTGTGTCT 60  
 GGGCTGTGTC TTTGCTGCCC CCAGACAGCG CCAGTCCACC CTGTGTGTCT TCCTGTGTGTA 120  
 CTTGAGAAC AATCTACACG ACAGGACAGT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180  
 CACTCGGGTG GCAGAGATGC GTGGAGATGC GAAATCTCTG GGGCCTGCGC TGCTGTCTCT 240  
 CCAGAGGCAA CTGTCCCTGC CGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300  
 GCGAACCCCA CGGTACGGGG TCCAGACCTT GGCAGATTC CAAACCTTTG AGGGCGACCT 360  
 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAC TACTCGAAG ACTTGCCTGC 420  
 GGCCTGTATT GACGAGCGCT TTGCGCGCGC CTTCGCACTG TGGAGCGCGG TGACGCGCGT 480  
 CACCTTCACT CGCGTGTACA GCGGGAGCGC AGACATGCTC ATCCAGTTTG GTGTGCGGGA 540  
 GCACGGAGAC GGGTATCCCT TOGACGGGAA GGACGGGCTC CTGGCACAAG CCTTTCCTCC 600  
 TGGCCCGGCG ATTCAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCTGGGCAA 660  
 GGGCGTCTGT GTTCCAACTC GGTTTGGAAA CGCAGATGGC GCGGCGTGGC ACTTCCCTCT 720  
 CATCTTCGAG GCGCGCTCCT ACTCTGCTCT CACCAACGAC GGTGCTCTCG ACGGCTTGGC 780  
 CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTGGCTTCTT GCCCAGCGA 840  
 GAGACTCTAC ACCCGGACG GCAATGTCTG TGGGAAACCC TGCCAGTTTC CATTCATCTT 900  
 CCAAGGCCAA TCCTACTCGG CTGCAACAC GACGCTCGC TCCGACGGCT ACCGCTGGT 960  
 GCGCACCAAC GCCAACTACG ACCGGGACAA GCTCTTGGC TTCTGCGCGA CCCGAGCTGA 1020  
 CTCGACGGTG ATGGGGGGCA ACTCGGCGGG GAGCTGTGCT GTCTTCCCTT TCACTTTCCT 1080  
 GGGTAAGGAG TACTCGACCT GTACCAGCGA GGGCGCGCGA GATGGGCGCC TCTGGTGTGC 1140  
 TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200  
 TTTGTTCCTC TTGGCGCGCT ATGAGTTCGG CCACGCGCTG GGCCTAGATC ATTCTCAGT 1260  
 GCGGAGGCGC CTATGTATCC CTATGTACCG CTTCAGTAG GGGCCCCCTT TGCTAAGGA 1320  
 CGACGTGAAT GGCATCGCGC ACCTCTATGG TCTCGCCCTT GAACCTGAGC CACGGCTCTC 1380  
 AACCAACACC ACACCGCGAG CCACGCTTCC CCGACCGGTC TGCCCCACCG GACCCCCAC 1440  
 TGTCAACCCC CTAGAGCGCC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC 1500  
 AGGTCCCCCG ACTGCTGGCC CTCTACGGC CACTACTGTG CCTTTGAGTC CGGTGGACGA 1560  
 TGCTTGCAAC GTGAACATCT TCAGACCCAT CGCGAGGATT GGGAAACAGC TGTATTGTGT 1620  
 CAAGGATGGG AAGTACTGCG GATTCTCTGA GGGCAGGGGG AGCGCGCGCG AGGGCCCCCT 1680  
 CCTTATCGCC GACAGTGGC CGCGCTGCC CGCAGAGCTG GACTCGGTCT TTAGGAGGCC 1740  
 GCTCTCCAGG AAGCTTTTCT TCTCTCTGCG GCGCGAGGTG TGGGTGTACA CAGGCGGCTC 1800  
 GGTGTCTGGC CGGAGCGGCT TGGACAAGCT GGGCCTGGGA GCGGACGCTG CCCAGGTGAC 1860  
 CGGGGCCCCC CGGAGTGGCA GGGGGAAGAT GCTGCTGTTC AGCGGGCGCG CCCTCTGGAG 1920  
 GTTCGAGCTG AAGGCGCAGA TGGTGGATCC CGGAGCGGCC AGCGAGGTGG ACGGATGTT 1980  
 CCGCGGGGTG CCTTTGGACA CGCAGCAGT CTTCAGTAC CGAGAGAAAG CCTATTCTG 2040  
 CCAGGACCGC TTCTACTGGC GCGTGTGTTT CCGGAGTGGG TTGAACAGG TGGACCAAGT 2100  
 GGGCTACGTG ACCTATGACA TCCTGCASTG CCCTGAGGAC TAGGGCTCCG GTCTGCTCTT 2160  
 GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCGGATA 2220  
 CAAACTGTGA TCTGTTCTG GAGGAAAGGG AGGAGTGGAG GTGGCTGGG CCCTCTCTTC 2280  
 TCACCTTTGT TTTTGTGTG AGTGTCTCTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 126 Protein sequence:  
 Protein Accession #: NP\_004985.1

1 11 21 31 41 51  
 MSLWQPLVLV LVLGCGCFAP PRQRQSTLVL FPGDLRTNLT DRQLAEELY RYGYTRVAEM 60  
 RGESKSLGPA LLLQKQLSL PETGELDSAT LKAMRTPRCG VPDLGRFQTF EGDLLKWHEN 120  
 ITYWTQNYSE DLEPRAVIDDA FARAPALWSA VTPLTFTRVY SRDADIVIOF GVAEHGDDYP 180  
 PDGKDGLLAH AFPFGPIQG DAHPDDELW SLGKGVVVPF RFGNADGAAC HPFFIFEGRS 240  
 YSACTTDRG DGLPWCSTTA NYDTDDRFGP CPSERLYTRD GNADGKPCQF PFIFQGSYS 300  
 ACTTDRGSDG YRWCAATTANY DRDKLFGFCP TRADSTVMGG NSAGELCVFF PTFLGKEYST 360  
 CTSEGRGDGR LWCATTNFD SDKKWGCPD QGYSLFLVAA HEPGHALGLD HSSVPEALMY 420  
 PMYRTEBPP LKEDDVNGIR HLYGPRPEPE PRPPTTTTPQ PTAPPTVCPT GPPTVHPSER 480  
 PTAGTGPPTS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFDGKYW 540  
 RFSEGRGSRP QGPFLIADKW PALPRKLDV FZEPLSKLP FPGSRQVWVY TGASVLPGR 600  
 LDKLGLGADV AQVTGALRSR RGKMLLFSGR RLWRPVDVQA MVDPRSASEV DRMPFPGVPLD 660  
 THDVFQYREK AYPQDRFYW RVSSRSELNQ VDQVGYVTYD ILQCPED

Seq ID NO: 127 DNA sequence  
 Nucleic Acid Accession #: NM\_004181  
 Coding sequence: 32-670

1 11 21 31 41 51  
 GCAGAAATAG CCTAGGGAGA TCAACCCCGA GATGCTGAAC AAAGTGTCTG CCGGCTGGG 60  
 GGTGCGCGGC CAGTGGCGCT TCGTGGACGT GCTGGGCTG GAAGAGGAGT CTCTGGGCTC 120  
 GGTGCCAGCG CCGTCTGCGC CGCTGCTGCT GCTGTTTCCC CTCACGGCCC AGCATGAGAA 180  
 CTTAGGAAA AAGCAGATTG AAGAGCTGAA GGGACAAGAA GTTAGTCTTA AAGTGTACTT 240  
 CATGAAGCAG ACCATTGGGA ATTCTGTGG CACAATCGGA CTTATTCAAG CAGTGGCCAA 300  
 TAATCAAGAC AAACCTGGAT TTGAGGATGG ATCAGTTCTG AAACAGTTTC TTTCTGAAAC 360  
 AGAGAAAATG TCCCTGAAG ACAGAGCAAA ATGCTTTGAA AAGAAATGAG CCATACAGGC 420  
 AGCCCATGAT GCGGTGGCAC AGGAAGGCCA ATGTCGGGTA GATGACAAGG TGAATTTCAC 480  
 TTTTATTCTG TTAAACAACG TGGATGGCCA CCTCTATGAA CTGTATGAC GAATGCCTTT 540  
 TCCGCTGAAC CATGCGGCCA GTTCAGAGGA CACCTGCTG AAGGACGCTG CCAAGGTGTG 600  
 CAGAGAAATC ACCGAGCGTG AGCAAGGAGA AGTCCGCTTC TCTCGCGTGG CTCTCTGCAA 660  
 GGCAGCCTAA TGCTCTGTGG GAGGGAATTT GCTGATTCCC CCTCTTCCCT TCAACATGAA 720  
 AATATATACC CCCCATGAG TCTAAAATGC TTCAGTACTT GTGAACACA GCTGTCTTTC 780  
 TGTCTGCGAG ACACGCTTTC CCCTCAGCCA CACCCAGGCA CTTAAGCACA AGCAGAGTGC 840  
 ACAGCTGTCC ACTGGGCGAT TGTGGTGTGA GCTTCAGATG GTGAAGCATT CTCCCCAGTG 900  
 TATGTCTTGT TCTCGATATC TAACGCTTTA AATGGCTACT TTGTTTCTG TCTGTAAAGT 960  
 AAGACCTTGG ATGTGGTTAT GTTGTCTTAA AGAATAAATT TTGCTGATAG TAGC

Seq ID NO: 128 Protein sequence:  
 Protein Accession #: NP\_004172

1 11 21 31 41 51  
MLNKVLSRLG VAGQWRFDV LGLEESLGS VPAPACALLL LPFLTAQHEN FRKKQIEELK 60  
GQEVSPKVYP MKQTTIGNSCG TIGLIHAVAN NQDKLGFEDG SVLKQFLSET EKMSPEDRAK 120  
CFEKNEALQA AHDAVAEGGQ CRVDDKVNPH FIFLNNVDGH LYELDGRMPF PVNHGASSED 180  
TLKDAARKVC REFTEREQGE VRFSAVALCK AA

Seq ID NO: 129 DNA sequence  
Nucleic Acid Accession #: NM\_000213  
Coding sequence: 127-5385

1 11 21 31 41 51  
CGCCGCGCGG CTGCAGCCCC ATCTCCTAGC GGCAGCCAG GCGCGGAGGG AGCGAGTCCG 60  
CCCCAGAGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GGCTGGCCGG GAGAGGGAGG 120  
AAGAGGATGG CAGGGCCACG CCCCAGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC 180  
AGCGTCAGCC TCTCTGGGAC CTTGGCAAAC CGCTGCAAGA AGGCCCCAGT GAAGAGCTGC 240  
ACGGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACAGAT GTTCAGGGAC 300  
CGCGCTGCA ACACCCAGGC GGAGCTGCTG GCGCGGGCT GCCAGGGGA GAGCATCGTG 360  
GTGATGGAGA CAGGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCAC CCGCGGCGCG 420  
AGCCAGATGT CCCCACAGG CCGTGGGGTC CGCTGCGGCG CCGGTGAGGA GCGGCATTTT 480  
GAGCTGGAGG TGTCTTGAGCC ACTGGAGAGC CCGGTGAGCC TGTACATCCT CATGGACTTC 540  
TCCAACCTCA TGTCCGATGA TCTGGACAAC CTCAAGAAGA TGGGGCAGAA CCGTGCCTCG 600  
GTCTTGAGCC AGCTCACCAG CGACTACACT ATTGGATTGT GCAAGTTTGT GGACAAAGTC 660  
AGCGTCCGCG AGACCGGACAT GAGGCGTGAG AAGCTGAAGG AGCCCTGGCC CAACAGTGAC 720  
CCCCCTTCT CTTCAAGAA CGTCAACAGC CTGACAGAA AGTGTGATGA GTTCGGGAAT 780  
AAATGACAG GAGAGCGGAT CTCAGGCAAC CTGAGTGCTC CTGAGGGCGG CTTGATGCC 840  
ATCCTGCAGA CAGCTGTGTG CAGCAGGAGC ATTGGCTGGC GCGCGGACAG CACCACCTG 900  
CTGGTCTTCT CCACCGAGTC AGCCTTCCAC TATGAGGCTG ATGGCGCCAA CGTGTGGCT 960  
GGCATCATGA GCGCAAGCA TGAACGGTGC CACCTGGACA CCAAGGGCAC CTACACCCAG 1020  
TACAGGACAC AGGACTACCC GTCCGTGCCC ACCCTGGTGC GCCTGCTGCG CAAGCACAAC 1080  
ATCATCCCA TCTTGTCTGT CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC 1140  
TATTTCCCTG TCTCTCACT GGGGTGCTG CAGGAGGACT CGTCAACAT CGTGAGCTG 1200  
CTGGAGGAG CTTTCAATCG GATCCGCTCC AACCTGGACA TCGGGCCCT AGACAGCCCC 1260  
CGAGGCTTTC GGACAGAGGT CACCTCCAAG ATGTTCAGAA AGACGAGGAC TGGTCCCTTT 1320  
CACATCCGCG GGGGGGAGT GGGTATATAC CAGGTGACGC TCGGGCCCT TGAGCAGTG 1380  
GATGGGAGCG ACCTGTGCGA GCTGCCGAG GACCAAGAG GCAACATCA TCTGAACTC 1440  
TCCTTCTCCG ACAGCTCAA GATGAGCGCG GGCATCATCT GTGATGTGTG CACCTGCGAG 1500  
CTGCAAAAG AGGTGCGGTC AGCTCGCTGC AGCTTCAACG GAGACTTGT GTGCGGACAG 1560  
TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620  
GACATTACGC CCTGCTGCG GGAGGGCGAG GACAAGCCGT GCTCGGCGCG TGGGAGTGC 1680  
CAGTGGCGCG ACTGTGTGTG CTACGGCGAA GCGCGTACG AGGGTCAGTT CTGCGAGTAT 1740  
GACAACTTCC AGTGTCCCGC CACTTCCGGG TTCTCTGCA ATGACCGAGG ACCTGTCTCC 1800  
ATGGGCCAGT GTGTGTGTA GCTGTGTG GCGTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCTCAGC 1860  
AATGCCACCT GCATCGACAG CAATGGGGGC ATCTGTAATG GAGGTGGCCA CTGTGAGTGT 1920  
GGCCGCTGCC ACTGCCACCA GCAGTGGCTC TACAGGACA CCACTGCGGA GATCAACTAC 1980  
TCGGGATACC AGCCGGGCT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040  
GGCAGCGGCG AGAAGAGAGG GCGCAGTGT GAGGAATGCA ACTTCAAGGT CAAGATGTTG 2100  
GACGAGCTTA AGAGAGCCGA GGAGGTGTGT GTGCGCTGCT CTTCCGCGGA CGAGGATGAC 2160  
GACTGCACCT ACAGCTACAC CATGGAAGGT GACGGCGCCC CTGGGCCCAA CAGCACTGTC 2220  
CTGGTGACA AGAAGAGAGA CTGCCCTCCG GGTCTCTTCT GGTGGCTCAT CCCCCTGCTC 2280  
CTCTCTCTCC TGCGCTCTCT GCGCTGCTA CTGCTGCTAT GCTGGAAGTA CTGTGCTGTC 2340  
TGCAAGGCTC CGCTGGCAT TCTCCGTGTC TGCAACGAG GTACATGTT GGGCTTTAAG 2400  
GAAGACCACT ACATGCTGG GGAGAACCTG ATGGCTCTG ACCACTTGA CAGCCCATG 2460  
CTGCGCAGCG GGAACCTCAA GGGCGGTGAC GTGGTCCGT GGAAGTCAAC CAACAACATG 2520  
CAGCGGCTTG GCTTTGCCAC TCATGCCGCG ASCATCAACC CCACAGAGCT GGTGCCCTAC 2580  
GGCTGTCTCT TGCGCTTGG CCGCTTTGCG ACCGAGAACC TGCTGAAGCC TGACACTCGG 2640  
GAGTGGCGCC AGCTGGGCCA GGAGGTGGAG GAGAACTGA ACGAGGTCTA CAGGAGATC 2700  
TCGGGTGATC ACAAGCTCCA GCAGACCAAG TTCCGCGAGC AGCCCAATGC CGGGAAGAG 2760  
CAGAGCCACA CCAATTGTGA CACAGTGTG ATGGCGCCCC GCTCGGCCAA GCGCGCCCTG 2820  
CTGAAGCTTA CAGAGAAGCA GGTGAACAG AGGGCTTCC ACGACCTCAA GGTGGCCCCC 2880  
GGCTACTACA CCTCACTGCG AGACAGGAC GCCCGGGCA TGGTGGAGTT CCAGGAGGGC 2940  
GTGGAGCTGG TGGAGTACG GGTGCCCTC TTTATCCGCG CTGAGGATGA CAGCAGAGAG 3000  
CAGCTGCTGG TGGAGGCCAT CGACGTGCCC GCAGGCACTG CCACCTCGG CCGCGGCTG 3060  
GTAAACATCA CCAATCATCA GAGCAAGCC AGAGAGCTGG TGTCTTTGA GCAGCTGAG 3120  
TTCTCGTCA GCGCGGGGA CAGGTGGCC CGCATCCCTG TCATCCGGCG TGTCTGGAC 3180  
GGCGGGAAGT CCCAGTCTC CTACCGACA CAGGATGGCA CCGCGCAGGG CAACCGGGAC 3240  
TACATCCCCG TGGAGGTGA GCTGCTGTC CAGCTGGGG AGGCTTGAAG AGAGCTGACG 3300  
GTGAAGCTCC TGGAGCTGCA AGAAGTTGAC TCCCTCTGCG GGGCGCGCA GGTCCGCGT 3360  
TTCCAGTCC AGCTCAGCAA CCTTAAGTTT GGGGCCACC TGGGCCAGCC CCACTCCACC 3420  
ACCATCATCA TCAGGGACCC AGATGAACG GACCGGAGCT TCACGAGTCA GATGTTGTCA 3480  
TCACAGCCAC CCCCTCAGCG CGACCTGGCG GCCCGCAGA ACCCAATGC TAAGGCCGCT 3540  
GGGTCCAGGA AGATCCATT CAACCTGGCT CCCCCTTCTG GCAAGCCAAT GGGGTACAGG 3600  
GTAAAGTACT GGATTACGGG TGACTCCGAA TCCGAAGCCC ACCTGTCTGA CAGCAAGGTG 3660  
CCCTCAGTGG AGCTCACCAA CCTGTACCG TATTGCGACT ATGAGATGAA GGTGTGCGCC 3720  
TACGGGGCTC AGGGCGAGGG ACCCTACAGC TCCCTGGTGT CCGTGGCGAC CCACAGGAA 3780  
GTGCGCAGCG AGCCAGGGCG TCTGGCCTTC AATGTGCTCT CCTCCAGGT GACCCAGCTG 3840  
AGCTGGGCTG AGCGGGCTGA GACCAACGCT GAGATCACAG CCTACGAGGT CTGCTATGGC 3900  
CTGGTCAAGC ATGACAACCG ACCTATTGGG CCCATGAAGA AAGTGTGTTG TGACAACTCT 3960  
AAGAACCGGA TGTGCTTAT TGAGAACCTT CCGGAGTCCC AGCCCTACCG CTACAGGTTG 4020  
AAGGCGGCA ACAGGGGCGG CTGGGGGCTG GAGCGGAGG CCATCATCAA CCGTGGCCAC 4080  
CAGCCCAAGA GGCCTATGTC CATCCCATC ATCCCTGACA TCCCTATGTT GGACGCCCCG 4140  
AGCGGGGAGG ACTACGACG CTTCTTATG TACAGGATG ACCTTCTAG CTCTCCATCG 4200  
GGCAGCCAGA GGCCTACGCT TCCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAC 4260  
TTTGCTTCC CCGGCAGCAG CAACTCCCTG CACAGGATGA CCACGACAG TGCTGCTGCG 4320  
TATGGCACCC ACCTGAGCCC ACAAGTGGCC CACCGGCTGC TAAGCACATC CTCACCCCTC 4380

25 Seq ID NO: 130 Protein sequence:  
Protein Accession #: NP\_000204

Seq ID NO: 131 DNA sequence  
Nucleic Acid Accession #: BC004372  
Coding sequence: 132..2231

238

## WO 02/086443

PCT/US02/12476

CCCAATCGAC AACAGGGACA GCTGCGCCT CAGCTCATAC CAGCCATCCA ATGCAAGGAA 1320  
GGACAACACC AAGCCACAGAG GACAGTTCTT GGACTGATTT CTTCACCCCA ATCTCACACC 1380  
CCATGGGACG AGGTGATCAA GCAGGAAGAA GGATGGATAT GGACTCCAGT CATAGTACAA 1440  
CGCTTCAGCC TACTGCAAAAT CCAACACAG GTTTGGTGGA AGATTGGAC AGGACAGGAC 1500  
CTCTTTCAAT GACAAAGCAG CAGAGTAATT CTCAGAGCTT CTCTACATCA CATGAAGGCT 1560  
TGGAGAAGA TAAAGACCAT CCAACAACCTT CTACTCTGAC ATCAAGCAAT AGGAATGATG 1620  
TCACAGGTGG AAGAAGAGAC CCAATCATT CTGAAGGCTC AACTACTTTA CTGGAAGGTT 1680  
ATACCTCTCA TTACCCACAC ACGAAGGAAA GCAGGACCTT CATCCAGTG ACCTCAGCTA 1740  
AGACTGGGTC CTTTGGAGTT ACTGCAGTTA CTGTTGGAGA TTCCAACTCT AATGTCAATC 1800  
GTTCTTATC AGGAGACCAA GACACATTCC ACCCCAGTGG GGGTCCCAT ACCACTCATG 1860  
GATCTGAATC AGATGGACAC TCACATGGGA GTCAAGAAGG TGGAGCAAAC ACAACCTCTG 1920  
TGCTATAAG GACACCCCAA ATTCCAGAAAT GGCTGATCAT CTGCGCATCC CTCTGGCCT 1980  
TGGCTTTGAT TCTTGCAATT TGCATTGCAG TCAACAGTCG AAGAAGGTGT GGGCAGAAGA 2040  
AAAAGCTAGT GATCAACAGT GGCATGGAG CTGTGGAGGA CAGAAAGCCA AGTGGACTCA 2100  
ACGGAGAGGC CAGCAAGTCT CAGGAAATGG TGCATTGGT GAACAAGGAG TCGTCAGAAA 2160  
CTCCAGACCA GTTTATGACA GCTGATGAGA CAAGGAACCT GCAGATGTG GACATGAAGA 2220  
TTGGGGTGTG ACACCTACAC CATTATCTTG GAAAGAAACA ACCGTGGAA ACATAACCAT 2280  
TACAGGGAGC TGGGACACAT AACAGATGCA ATGTGCTACT GATTGTTTCA TTGCGAATCT 2340  
TTTTTAGCAT AAAATTTTCT ACTCTTAAAA AAAAAAAAAA AAAAAA

Seq ID NO: 132 Protein sequence:  
Protein Accession #: AAH04372

1 11 21 31 41 51  
MDKFWHAAM GLCLVPLSLA QIDLNITCRF AGVPHVEKNG RYISIRTEAA DLCKAFNSTL 60  
PTMAQMEKAL SIGFETCRYG FIEGHVVIPIR IHPNSICAA NTVGVILTSN TSQYDTCVFN 120  
ASAPFEEDCT SVTDLPNAPD GPITITIVNR DGTTRYVQKE YRTNPEDIYP SNPTDDVSS 180  
GSSSERSSTS GGYIYFTTSP VHPIDEDSP WITDSTRIP ATSTSSNTIS AGWEPNEENE 240  
DERDRHLSFS GSGIDDDSDP ISSTISTTFR AFDTKQND WTQWNPSSHN PEVLQTTTR 300  
MTDVRNGTTF AYSGWNPNPEA HPPLIHHEHH EEEETPHSTS TIQATPSSTT EETATQKEQW 360  
FGNRWEGYR QTPREDSHST TGTAASAH TSEFMQGRTP SPEDSSWTF FNPISHPMGR 420  
GHQAGRRMDM DSSHSTTLQF TANPNTGLVE DLDRTGPLSM TTQSSNSQSF STSHGLEED 480  
KDHPTTSLT SGNRNDVTGG RRDPNRSEGS TTLLEGYTS YPHTKESRTF IPVTSAKTGS 540  
FGVAVTVGD SGNVNRSLG GDQTFPHPSG GSHHTHSGES DGHSHSGSEG GANTTSGPIR 600  
TPQIPWLI LALILALALI LAVCIAVNSR RRCGQKKLV INSGNGAVED RKPSGLNGEA 660  
SKSQEMVHLV NKESSETFDQ FMTADETRNL QNVDMKIGV

Seq ID NO: 133 DNA sequence  
Nucleic Acid Accession #: NM\_002882  
Coding sequence: 150-755

1 11 21 31 41 51  
CGAGGTTCCG GTGCTGGGGC GGAGGGAAGA GCGGGCGGGC GGGAGGCGCC GCGGCCAGAC 60  
GCGGAGGAA GGAGCTACGA GTAGCCGCGC AGAGGCGCGC GAGCCAGCGA CGACCGACCC 120  
AGCCGAGCG CGCCGCGCGC CGCGGCCCA TGGCGGCGC CAAGGACACT CATGAGGACC 180  
ATGATACTTC CACTGAGAA ACAGACGAGT CCAACCATGA CCTCAGTTT GAGCCAATAG 240  
TTTCTCTTCC TGAGCAAGAA ATTAAACAC TGAAGAAGA TGAAGAGGAA CTTTTTAAAA 300  
TGGGGCAAAA ACTGTTCCGA TTTGCCCTG AGAAGCATCT CCCAGAATGG AAGGAGCGAG 360  
GCACTGGTGA CGTCAAGCTC CTGAAGCACA AGGAGAAAGG GGCCATCCGC CTCTCTCATG 420  
GGAGGACAAA AGCCCTGAAG ATCTGTGCCA ACCACTACAT CACGCGGATG ATGGAGCTGA 480  
AGCCCAAGCG AGGTAGCGAG CGTGCTGGG TCTGGAACAC CCAAGCTGAC TTCGCCGAGC 540  
AGTGCCCAAA GCCAGAGCTG CTGGCCATCC GCTTCTGAA TGCTGAGAA GCACAGAAAT 600  
TCAAAACAAA GTTTGAAGAA TGCAGGAAAG AGATCGAAGA GAGAGAAAG AAAGCAGGAT 660  
CAGGCAAAAA TGATCATGCC GAAAAGTGG CGGAAAGCT AGAAGCTCTC TCGGTGAAGG 720  
AGGAGACCAA GGAGGATGCT GAGGAGAAGC AATAAATCGT CTTATTATAT TTCTTTTTC 780  
TCTCTTCTCT TTCTTTT TAAAAAATT TACCCTGCCC CTCTTTTTCG GTTTGTTTTT 840  
ATTCTTTTCT TTTTACAAGG GAGCTTATAT AAAGAACTGA ACTC

Seq ID NO: 134 Protein sequence:  
Protein Accession #: NP\_002873

1 11 21 31 41 51  
MAAAKDTHEH DDTSTENTDE SNHDPQFEPI VSLPEQEIKT LSEDEEELPK MRALKFRPAS 60  
ENDLPEWKER GTGDKLLKH KEKGAIKLLM RRDKTLKICA NHYITPMEL KFNAGSDRAW 120  
VWNTHADPAD ECPKPELLAI RFLNAENAK FTKFRECRK EIEEREKKAG SGKNDHAEKV 180  
AEKLEALS VK EETKEDAEK Q

Seq ID NO: 135 DNA sequence  
Nucleic Acid Accession #: NM\_000077.2  
Coding sequence: 277-742

1 11 21 31 41 51  
CCCAACCTGG GCGGACTTCA GGTGTGCCAC ATTGCTAAG TGCTCGGAGT TAATAGCACC 60  
TCTCTCGAGC ACTGCTCAC GGCCTCCCT TGCTGGGAAA GATACGCGG TCCCTCCAGA 120  
GGATTGAGG GACAGGGTGC GAGGGGGCTC TTCCGCCAGC ACCGAGGAA GAAAGAGGAG 180  
GGGCTGGCTG GTCAACAGAG GGTGGGGGGC ACCGCGTGC CTGCGGGCTC GCGGAGAGGG 240  
GGAGAGCAGG CAGCGGGGGC CGGGAGCAG CATGGAGCCG GCGCGGGGA GCAGCATGGA 300  
GCCTTCGCT GACTGCTGG CCACGGCCGC GCGCGGGGT CCGGTAGAGG AGGTGCGGGC 360  
GCTGCTGGAG GCGGGGGGCG TGCCCAAGC ACCGAATAGT TACGCTCGGA GGCCGATCCA 420  
GGTCATGATG ATGGGCAGCG CCGAGTGCGC GGAGCTGCTG CTGCTCCAGC GCGCGGAGCC 480

CAACTGCGCC GACCCCGCCA CTCTCACCG ACCCGTGAC GACGCTGCC GGGAGGGCTT 540  
CCTGACACG CTGGTGGTGC TGCACCGGCG CGGGGCGCG CTGGACGTGC GCGATGCCTG 600  
GGGCGGCTG CCGGTGGACC TGGCTGAGGA GCTGGGCCAT CGCATGTGCG CACGGTACCT 660  
GCGCGGGCT GCGGGGGGCA CCAGAGGCG TAACCATGCC CGCATAGATG CCGCGGAAGG 720  
TCCTTCAGAC ATCCCGGATT GAAAGAACCA GAGAGGGTCT GAGAAACCTC GGGAAACTTA 780  
GATCATCAGT CACCGAAGGT CCTACAGGGC CACAACCTGC CCGCCACAA CCCACCCCGC 840  
TTTCGTAGTT TTCAATTAGA AAATAGAGCT TTTAAAAATG TCCTGCCTTT TAACTAGAT 900  
ATATGCTTTC CCCCCTACC GTAAATGTCC ATTTATATCA TTTTATATAT ATTCTTATA 960  
AAATGTAAAA AAGAAAAACA CCGCTTCTGC CTTTCTACTG TGTGGAGTT TTCTGGAGTG 1020  
AGCATTACG CCTAAGCGC ACATTTCATGT GGGCATTTC TGGAGGCTC GCAGCCTCG 1080  
GAAGCTGTG ACTTCATGAC AAGCATTTTG TGAAGTAGGG AAGCTCAGGG GGGTTACTGG 1140  
CTTCTCTGA GTCACTGCT TAGCAATGG CAGAACCAAA GCTCAATAA AAATAAAATA 1200  
ATTTTCATTC ATTCACTC

Seq ID NO: 136 Protein sequence:  
Protein Accession #: NP\_000068.1

1 11 21 31 41 51  
MEPAAGSSME PSADWLATAA ARGVVEEVR LLEAGALPNA FNSYGRRIQ VMMSGARVA 60  
ELLLLHGAEP NCADPATLTR PVHDAAREGP LDTLVVLRH GARLDVRDAW GRLPVDLAEZ 120  
LGRDVARYL RAAAGGTRGS NHARIDAAEG PSDIPD

Seq ID NO: 137 DNA sequence  
Nucleic Acid Accession #: NM\_058196.1  
Coding sequence: 104-421

1 11 21 31 41 51  
TGTTGGGGG TCTGCTGGC GGTGAGGGG CTCTACACAA GCTTCCTTTC CGTCATGCCG 60  
GCCCCACCC TGGCTCTGAC CATTCGTTC TCTCTGGCAG GTCATGATGA TGGGCAGCGC 120  
CCGAGTGGG GAGCTGTGCG TGCTCCACGG CCGCGAGCCC AACTGCGCG ACCCCGCCAC 180  
TCTCACCCGA CCGGTGCAAG CGCTGCGCG GGAGGGCTTC CTGGACACGC TGGTGTGTCT 240  
GCACCGGGCC GGGGCGCGC TGGAGGTGCG CGATGCCTGG GCGCGCTGCG CCGTGSACCT 300  
GGCTGAGGAG CTGGGCCATC CGCATGTGCG ACGGTACCTG CCGCGGCTG CCGGGGGCAC 360  
CAGAGGCGT AACCATGCCC GCATAGATGC CGCGAAGGT CCCTCAGACA TCCCGATTG 420  
AAAGAACCA AGAGGCTCTG AGAAACCTCG GGAACCTTAG ATCATCAGTC ACCGAAGGTC 480  
CTACAGGGCC ACAACTGCC CCGCCACAAC CCACCCGCT TCGTAGTTT TCATTAGAA 540  
AATAGAGCTT TTAATAATGT CCGCTCTTTT AACGTAGATA TAAGCCTTCC CCCACTACCG 600  
TAAATGTCCA TTTTATCAT TTTTATATA TTTTATAAAA AATGTAAAAA AGAAAAACAC 660  
CGCTTCTGCC TTTTCACTGT GTTGGAGTTT TCTGGAGTGA GCATCAACGC CTAAGCGCA 720  
CATTCATGT GGCATTTCTT GCGAGCCTCG CAGCCTCCGG AAGCTGTGCA CTTTATGACA 780  
AGCATTTTGT GAACTAGGGA AGCTCAGGGG GGTACTGGC TTCTCTGAG TCACACTGCT 840  
AGCAATGGC AGAACCAAG CTCAAATAAA AATAAAATAA TTTTCATTCA TTCCTC

Seq ID NO: 138 Protein sequence:  
Protein Accession #: NP\_478103.1

1 11 21 31 41 51  
MMMSGARVAE LLLHGAEPN CADPATLTR PVHDAAREGP DTLVVLRHAG ARLDVRDAWG 60  
RLPVDLAEEL GHRDVARYL RAAAGGTRGS HARIDAAEG SDIPD

Seq ID NO: 139 DNA sequence  
Nucleic Acid Accession #: NM\_058197.1  
Coding sequence: 272-684

1 11 21 31 41 51  
CCCAACTGG GGGCACTTCA GGTGTGCCAC ATTGCTAAG TGCTCGGAGT TAATAGCACC 60  
TCCTCCGAGC ACTGCTCAC GCGTCCCTT TGCTGGAAA GATACCGCG TCCCTCCAGA 120  
GGAITTAGG GACAGGTCG GAGGGGGCTC TTCGCCAGC ACOGGAGGAA GAAAGAGGAG 180  
GGGCTGGCTG GTACACAGAG GGTGGGGCGG ACCGCTGCG CTGCGCGCT CCGGAGAGGG 240  
GGAGAGCAG CAGCGGGCGG CCGGGAGCAG CATGAGCGG CCGCGGGGA GCAGCATGGA 300  
GCGGCGGCG GGGAGCAGCA TGGAGCCTTC GGCTGACTGG CTGGCCACGG CCGCGGCCCG 360  
GGGTGCGGTA GAGGAGGTGC GGGCGCTGCT GGAGGCGGG GCGCTGCCA ACGCACCGAA 420  
TAGTTACGGT CGGAGGCGGA TCCAGGTGGG TAGAAGTCT GCAGCGGAG CAGGGGATGG 480  
CGGGCGACTC TGGAGGACGA AGTTTGCAGG GGAATTGGAA TCAGTAGCG CTTGATTCT 540  
CCGGAATAAG GGGAGGCTTC CTGGGAGTT TTCAGAAGG GTTTGTAATC ACAGACCTCC 600  
TCCTGGCGAC GCGCTGGGG CTTGGGAAC CAAGGAAGAG GAATGAGGAG CCACGCGCT 660  
ACAGATCTCT CGAATGCTGA GAAGATCTGA AGGGGGGAAC ATATTGTAT TAGATGGAAG 720  
TCATGATGAT GGGCAGCGCC CGAGTGGCG AGCTGCTGCT GCTCCACGCG CCGGAGCCCA 780  
ACTGCGCGTA CCGCGCACT CTCACCGAC CCGTGACGA CGCTCCCGG GAGGGCTTCC 840  
TGACACGCT GGTGTGCTG CACCGGGCG GCGCGGCTC GGAAGTGGC GATGCTGGG 900  
GCCGTCTGCC CGTGGACCTG GCTGAGGAGC TGGCCATCG CGATGTGCA CGGTACCTGC 960  
GCGCGGCTGC GGGGGGCACC AGAGGCACTA ACCATGCCCG CATAGATGCC GCGGAAGGTC 1020  
CCTCAGACAT CCGGATTAGA AAGAACAGA GAGGCTCTGA GAAACCTCG GAACTTAGAT 1080  
CATCAGTCA CGAAGGTCTT ACAGGGCCAC AACTGCCCG GCCACAACCC ACCCGCTTT 1140  
CGTAGTTTTC ATTTAGAAAA TAGAGCTTTT AAAAAATGCC TGCTTTTAA CTTAGATATA 1200  
TGCTTCCCC CACTACCGTA AATGTCCATT TATATCATT TTTATATATT CTTATAAAA 1260  
TGTAAAAAG AAAAAACCG CTTCTGCTT TCACTGTGT TGGAGTTTC TGGAGTGAGC 1320  
ACTCAGCGCC TAAGCGACA TTCATGTGGG CATTTCTTC GAGCCTCGCA GCCTCCGGA 1380  
GCTGTGACT TCATGACAG CATTTTGTGA ACTAGGGAAG CTCAGGGGG TTAAGGCTT 1440  
CTCTGAGTC ACATGCTAG CAAATGGCAG AACCAAGCT CAAATAAAA TAAATAATT 1500

Seq ID NO: 140 Protein sequence:  
Protein Accession #: NP\_478104.1

1 11 21 31 41 51  
MEPAAGSSME PAAGSSMEPS ADWLATAAAR GRVEEVRLALL RAGALPNAPN SYGRRPIQVG 60  
RRSAAGAGDG GRLWRKTFAG ELESQSASIL RKKGRLPGEF SEGVNHRPP PGDALGAWET 120  
KEEE

Seq ID NO: 141 DNA sequence  
Nucleic Acid Accession #: NM\_058195.1  
Coding sequence: 163-684

1 11 21 31 41 51  
CCTCCCTACG GCGCGCTCCG GCAGCCCTTC CCGCGTGC GC AGGGCTCAGA GCCGTTCCGA 60  
GATCTTGGAG GTCCGGGTGG GAGTGGGGGT GGGGTGGGGG TGGGGGTGAA GGTGGGGGGC 120  
GGGCGCGCTC AGGGAAGGCG GGTGCGCGCC TCGGGGCGG AGATGGGCG GGGCGGGTGC 180  
GTGGGTCCCA GTCTGCATT AAGGGGGCAG GAGTGGCGCT GCTCACTCT GGTGCCAAG 240  
GGGCGCGCAG CCGCTGCCGA GCTCGGCCCT GGAGGCGGCG AGAACATGTT GCGCAGGTTT 300  
TTGGTGACCC TCGGATTGCG GCGCGCGTGC GCGCGCGCG GAGTGAGGGT TTTCGTGTTT 360  
CACATCCGCG GGCTCACGCG GAGTGGGCA GCGCCAGGGG CGCCCGCCGC TGTGGCCCTC 420  
GTGCTGATGC TACTGAGGAG CCAGCGTCTA GGGCAGCAGC CGCTTCTAG AAGACCAGGT 480  
CATGATGATG GGCAGCGGCC GAGTGGCGGA GCTGCTGCTG CTCCAGCGCG CGGAGCCCAA 540  
CTCGCGCGAC CCGGCCATCT TCACCGGACC CGTGACGAC GCTGCCCGGG AGGGCTTCTT 600  
GGACACGCTG GTGGTGTCTG ACCGGGCGCG GCGCGGCTG GACGTGCGCG ATGCTGGGG 660  
CGGTCTGCCC GTGACCTTGG CTGAGGAGCT GGGCCATGCG GATGTGCGAC GGTACCTGCG 720  
CGCGGCTGCG GGGGCGACCA GAGGCAGTAA CCATGCCCGC ATAGATGCCG CGGAAGGTCC 780  
CTCAGACATC CCGATTGAA AGAACAGAG AGGCTCTGAG AACCTCGGG AACCTTAGAT 840  
CATCAGTACG CGAAGGTCTT ACAGGGCCAC AACTGCCCGC GCCACAAACC ACCCGCTTT 900  
CGTAGTTTTT ATTAGAAA TAGAGCTTT AAAAATGTCC TGCCCTTTAA CGTAGATATA 960  
TGCCCTTCCC CACTACGTA AATGTCCATT TATATCATTT TTTATATATT CTTATAAAAA 1020  
TGTAATAAAG AAAAACACCG CTTCTGCTTT TTCCTGTGT TGGAGTTTTT TGGAGTGAGC 1080  
ACTCACGCCC TAAGCGCACA TTCTATGTTG CATTCTTGC GAGCCTCGCA GCCTCCGGAA 1140  
GCTGTGACT TCATGACAAG CATTCTTGA ACTAGGGAAG CTCAGGGGGG TTAAGTGGCT 1200  
CTCTTGAGTC ACACGTCTAG CAAATGGCAG AACCAAGCT CAAATAAAAA TAAATAATT 1260  
TTCATTCAIT CACTC

Seq ID NO: 142 Protein sequence:  
Protein Accession #: NP\_478102.1

1 11 21 31 41 51  
MGRGRCVGPS LQLRQEWRC SPLVPKGGAA AAELGPGGGE NMVRFLVTL RIRACGPPR 60  
VRVFFVHPR LTGEWAAPGA PAVALVLM LRSRLGQQP LPRRPGHDDG QRPSSGAAAA 120  
PRRGAQLRRP RHSHPTARR CPGLPGHAG GAAPGRGAAG RARCLGPSAR GPG

Seq ID NO: 143 DNA sequence  
Nucleic Acid Accession #: NM\_018131  
Coding sequence: 412..1107

1 11 21 31 41 51  
GAAATTGCAC ACTTAAAGAC ATCAGTGGAT GAAATCACA GTGGGAAAGG AAAGCTGACT 60  
GATAAGAGA GACAGAGACT TTTGGAGAAA ATTCGAGTCC TTGAGGCTGA GAAGGAGAAG 120  
AATGCTTATC AACTACAGA GAAGGACAAA GAAATACAGC GACTGAGAGA CCAACTGAAG 180  
GCCAGATATA GTACTACCGC ATTGCTTGAA CAGCTGGAAG AGACAAAGAG AGAAGGAGAA 240  
AGGAGGAGAG AGGTGTTGAA AGCCTTATCT GAAGAGAAAG ACCTATTGAA ACAACAGTTG 300  
TCTGCTGCAA CCTCAGCAAT TGCTGAACTT GAAAGCAAAA CCAATACACT CCGTTTATCA 360  
CAGACTGTGG CTCCAAACCTG CTTCAACTCA TCAATAAATA ATATTATGA AATGGAAATA 420  
CAGCTGAAAG ATGCTCTGGA GAAAAATCAG CAGTGGCTCG TGTATGATCA GCAGCGGGA 480  
GTCTATGATA AAGGACTTTT AGCAAGATC TTTGAGTTGG AAAAGAAAAC GGAACAGCT 540  
GCTCATTAC TCCACAGCA GACAAAAAAG CCTGAATCAG AAGGTTATCT TCAAGAGAG 600  
AAGCAGAAAT GTTACAAGA TCTCTTGGCA AGTGCAAAAA AAGATCTTGA GGTGAAACGA 660  
CAAAACATAA CTCAGCTGAG TTTTGAACCTG AGTGAATTTC GAAGAAAATA TGAAGAAACC 720  
CAAAAGAAAG TTCACAATTT AAATCAGCTG TTGTATTAC AAAGAAGGGC AGATGTGCAA 780  
CATCTGGAAG ATGATAGGCA TAAACAGAG AAGATACAAA AACTCAGGGA AGAAGATGAT 840  
ATTGCTAGGG GAAAACCTGA AGAAGAGAAG AAGAGATCCG AAGAGCTCTT ATCTCAGGTC 900  
CAGTCTCTTT ACACATCTCT GCTAAAGCAG CAAGAGAAC AAACAAGGGT AGCTCTGTTG 960  
GAACACAGA TGCAAGCATG TACTTTAGAC TTTGAAAATG AAAAATCTGA CCGTCAACAT 1020  
GTGCAGCATC AATTGCATGT AATTCTTAAG GAGCTCCGAA AAGCAAGAAA AAATAACACA 1080  
TTTGAATCC TTGAACAGC TTCATGAGTT TGCCATCACA GAGCCATTAG TCACCTTCCA 1140  
AGGAGAGACT GAAAACAGAG AAAAAGTTGC CGCTCACCA AAAAGTCCCA CTGCTGCACT 1200  
CAATGGAAGC CTGGTGGATG GTCCCAAGTG CAATATACAG TATCCAGCCA CTGAGCATCG 1260  
CGATCTGCTT GTCCATGTGG AATACTGTTT AAAGTAGCAA AATAAGTATT TGTTTTGATA 1320  
TTAAAGATT CAATCTGTA TTTTCTGTTA GCTTGTGGGC ATTTTGAATT ATATATTCTA 1380  
CATTTTGAT AAAACTGCCT ATCTACCTTT GACACTCCAG CATGCTAGTG AATCATGTAT 1440  
CTTTTAGCT GCTGTGCATT TCTCTTGGCA GTGATACCTC CCTGACATGG TTCATCATCA 1500  
GGCTGCAATG ACAGAAATGTG GTGAGCAGCG TCTACTGAGA TACTAACATT TTGCACTGTC 1560  
AAAACTCTG GTAGGAAAAA GATAGCTCAG GTTATTGCTA ATGGGTTAAT GCACCAGCAA 1620  
GCAAAATATT TTAGTTTTCG GGGGTTTGA AAAATCAAG ATAATTAAAC AAGGATCTTA 1680  
ACTGTGTTG CATTTTTAT CCAAGCACTT AGAAAACCTA CAATCTTAAT TTTGATGTC 1740  
ATTGTTAAGA GGTGGTGATA GATACTATTT TTTTTCATA TTGTATAGCG GTTATTAGAA 1800



5 AAGTTGGGGA TTTTCTTGAT CTTTATGCT GCTTACCATT GAAACTTAAC CCAGCTGTGT 1860  
TCCCAACTCT TGTTCTGCGC ACAGAAACAGT ATCTGTTTGA GGCATAATCT TAAGTGGCCA 1920  
CACACAATGT TTTCTCTTAT GTTATCTGSC AGTAAGCTGA ACTGAATTA CATTAGCACA 1980  
TTCNGCTTAG CTTAAATTGT TAAATAAATC TTTAATAAAC CCATGTAGCC CTCTCATTGT 2040  
ATTGACAGTA TTTTAGTTAT TTTTGGCATT CTTAAGAGCTG GGCATGTAA TGATCAGATC 2100  
TTTGTTTGTC TGAACAGGTA TTTTATACA TGCTTTTGT AAACCAAAAA CTTTAAATT 2160  
TCTTCAGTT TCTAACATG CTTACCAGTG GGTACTGTGA AATGAGAAAA GAATAAAATT 2220  
ATTTAATGTT TT

10 Seq ID NO: 144 Protein sequence:  
Protein Accession #: NP\_060601

15 1 11 21 31 41 51  
| | | | | |  
MEIQLKDALE KNQOWLVDYQ QREVVYKGLL AKIFELEKKT ETAHSLPQQ TKKPESEGYL 60  
QEEKQKCYND LLASAKKDLK VERQITQLS FELSEFRKY ESTQKEVHNL NQLYSQRRR 120  
20 DVQHLEDDRH KTEKIQLKRE ENDIARGKLE EEKKRSEELL SQVQSLYTSL LKQEQEQTRV 180  
ALLBQQMQAC TLDENEKLD RQHVQHQHVV ILKELRKARK NNTVGILETA S

25 Seq ID NO: 145 DNA sequence  
Nucleic Acid Accession #: NM\_001168  
Coding sequence: 50..478

30 1 11 21 31 41 51  
| | | | | |  
CGCCAGATT TGAATCGCGG GACCCGTTGG CAGAGGTGGC GCGCGCGGCA TGGGTGCCCC 60  
GACGTTGCCC CTGCGCTGGC AGCCCTTTCT CAAGGACCAC CGCATCTCTA CATTCAAGAA 120  
CTGGCCCTTC TTGGAGGGCT GCGCCTGCAC CCGGAGAGCG ATGGCGGAGG CTGGCTTCAT 180  
CCACTGCCCC ACTGAGAACG AGCCAGAGCTT GCGCCAGTGT TTCTTCTGCT TCAAGGAGCT 240  
GGAAGGCTGG GAGCCAGATG ACCAGCCCAT AGAGGAACAT AAAAAGCATT CGTCCGGTTG 300  
CGCTTCTCTT TCTGTCAAGA AGCAGTTTGA AGAATTAACC CTGTGTGAAT TTTTGAAACT 360  
35 GGACAGAGAA AGAGCCAAGA ACAAAATTGC AAAGGAACC AACATAAGA AGAAAGAATT 420  
TGAGGAAACT GCGAAGAAAG TGCGCCGTGC CATOGAGCAG CTGGCTGCCA TGGATTGAGG 480  
CCTCTGGCGG GAGCTGCCCT GTCCAGAGT GGTGCACCA CTTCAGGGT TTATTCCTTG 540  
GTGCCACCAAG CCTTCTCTGT GCGCCCTTAG CAATGTCTTA GGAAGGAGA TCAACATTTT 600  
40 CAAATTAGAT GTTCAACTG TGCTCTCTGT TTGTCTTGA AGTGGCACA GAGGTGCTTC 660  
TGCTCTGTGA GCGGGTGTCT CTGGTAACAG TGGCTGCTTC TCTCTCTCTC TCTCTTTTTT 720  
GGGGGCTCAT TTTTGTCTGT TTGATTCCCG GGTCTACCAG GTGAGAAGTG AGGAGGGAAG 780  
AAGGCAGTGT CCTTTTGTCT AGAGCTGACA GCTTTGTTCG CGTGGGCAGA GCCTTCACAA 840  
GTGAATGTGT CTGAACCTCA TGTGTGTGAG GCTGTACAG TCCTGAGTGT GGACTTGGCA 900  
GTGCGCTGTT GAATCTGAGC TGCAGGTTCC TTATCTGTCA CACCTGTGCC TCCTCAGAGG 960  
45 ACAGTTTTTT TGTGTGTGTG TTTTGTGTT TTTTTTTTT GGTAGATGCA TGACTTGTGT 1020  
GTGATGAGAG AATGGAGACA GAGTCCCTGG CTCCTCTACT GTTTAACAAC ATGGCTTTCT 1080  
TATTTTGTGT GAATGTGTAA TTCAAGAAAT AGCACAACAT ACAATTAAAA CTAAGCACA 1140  
AGCCATTCTA AGTCATTGGG GAAACGGGGT GAACTTCAGG TGATGAGGA GACAGAATAG 1200  
50 AGTATAGAGA AGCGTCTGGC AGATACTCCT TTTGCCACTG CTGTGTGATT AGACAGGCCC 1260  
AGTACGCGCG GCGGCACATG CTGGCGCTC CTCCTCAGA AAAAGGCAGT GGCCTAAATC 1320  
CTTTTAAAT GACTTGGCTC GATGCTGTGG GGGACTGGCT GGGCTGTGCG AGGCCGTGTG 1380  
TCTGTGAGCC CAACCTTCAC ATCTGTACAG TTCTCCACAC GGGGAGAGGA CGCAGTCCGC 1440  
CCAGGTCCCC GCTTCTTTTG GAGGCAGCAG CTCGGAAGTC GCGTAAAGAT 1500  
55 GATGGAATTG ATTGCGCCTC CTCCCTGTCA TAGAGCTGCA GGTGGATTG TTACAGCTTC 1560  
GCTGGAACCC TCTGAGGTC ATCTCGGCTG TTCTGAGAA ATAAAAAGCC TGTCAATTC

Seq ID NO: 146 Protein sequence:  
Protein Accession #: NP\_001159

60 1 11 21 31 41 51  
| | | | | |  
MGAPTLPFAW QPFLKDRIS TPKNWPFLG CACTPERMAE AGPIHCPTEN EPDLAQCFPC 60  
65 FKELEGWEPD DDPIEHKKH SSGCAPLSVK KQFEELTLGE FLKLDREKAK NKIAKETNNK 120  
KKEFEETAKK VRRATEQLAA MD

70 Seq ID NO: 147 DNA sequence  
Nucleic Acid Accession #: NM\_014176.1  
Coding sequence: 127-720

75 1 11 21 31 41 51  
| | | | | |  
GCGCGCAGCG CTGGTACCCC GTTGGTCCGC GCGTGTCTGC GTTGTGAGGG GTGTCACTC 60  
AGTGCATCCC AGGCAGCTCT TAGTGTGGAG CAGTGAACCT TGTGTGGTTC CTCTACTTGT 120  
GGGATCATGC AGAGAGCTTC ACCTCTGAAG AGAGAGCTGC ACATGTTAGC CACAGAGCCA 180  
CCCCCAGGCA TCACATGTGT GCAAGATAAA GACCAAAATG ATGACCTGCG AGCTCAAATA 240  
TTAGGTGGAG CCAACACACC TTATGAGAAA GGTGTTTTTA AGCTAGAAAT TATCAATCCT 300  
GAGAGGTACC CATTGGAACC TCCTCAGATC CGATTCTCA CTCCAATTTA TCATCCAAC 360  
80 ATTGATCTCT CTGGAAGGAT TTGTCTGGAT GTTCTCAAAT TGCCACCAAA AGGTGCTTGG 420  
AGACATCCG TCAACATGCG AACTGTGTGT ACCTCTATTC AGCTGCTCAT GTCAGAACCC 480  
AACCCTGATG ACCCGCTCAT GGCTGACATA TCCTCAGAAAT TTAATATATA TAAGCCAGCC 540  
TTCTCTAAGA ATGCCAGACA GTGGACAGAG AAGCATGCAA GACAGAAACA AAAGGCTGAT 600  
GAGGAAGAGA TGCTTGATTA TCTACAGAG GCTGTGACT CCAGAGTACA CAATCAACA 660  
85 GAGAAAGGA AGGCCAGTCA GCTAGTAGGC ATAGAAAAGA AATTTCAATC TGATGTTTGG 720  
GGGACTTGTG CTGGTTTCAT TTAGTTAATG TGTCTTTGCG CAGGTGATGC TAAGTTGCCT 780  
ACCTGAATT TTTTAAAA TATATTGAT GACATAAATT TTGTGTAGTT TATTTATCCT 840  
GTACATATGT ATTTTGAAT CTTTAAACC TGAAAAATA ATAGTCATTT AATGTTGAAA 900

WO 02/086443  
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

PCT/US02/12476

Seq ID NO: 148 Protein sequence:  
Protein Accession #: NP\_054895.1

```
1      11      21      31      41      51
MQRASRLKRE LHM LATEPPP GITCWQDKDQ MDDLRAQILG GANTPYEKGV FKLEVIIPER 60
YPFEPPIQIRF LTPYIHPNID SAGRICLDVL KLPPKGAWRP SLNIATVLTS IQLMSEPNP 120
DDPLMADISS EFYKNKPAFL KNARQWTEKH ARQKQKADSE EMLDNLPEAG DSRVHNSTQK 180
RKASQLVGLIE KKFHPDV
```

Seq ID NO: 149 DNA sequence  
Nucleic Acid Accession #: NM\_003812  
Coding sequence: 224-2722

```
1      11      21      31      41      51
TCCTCTGCGT CCCGCCCGCG GAGTGGCTGC GAGGCTAGGC GAGCCGGGAA AGGGGGGGCC 60
GCCAGCCCCG GAGCCCCGCG CCCCGTGGCC CGAGCCCGGA GCCCCTGCCC CGCGGCGGCA 120
CCATGCGCGC CGAGCGCGCG TGACCGGCTC CGCCCGCGGC CGCCCGCGAG CTAGCCCGGC 180
GCTCTGCGCG GCCACACGGA CGCGCGCGCG GGAGCTATGA GCCATGAAGC CGCCCGGCAG 240
CAGCTCGCGG CAGCGCGCCC TGGCGGGCTG CAGCCTTGCC GGCGCTTCTT CGCGGCCCCA 300
ACGCGGCCCC GCCGCTCGCG TGCCCTGCCAG CCGCCCGGCC CGCACGCGCC CCTGCCGCGT 360
GCTTCTCGTC CTCTCTCTGC TGCCCTCGCT CGCGCGCTCG TCCCGGCCCC GCGCTTGGGG 420
GGCTGCTGCG CCACGCGCTC CGCATTTGGA TGAAACTGCA GAAAAAATT TGGGAGTCCT 480
GGCAGATGAA GACAATACAT TGCAACAGAA TAGCAGCAGT AATATCAGTT ACAGCAATGC 540
AATGCAAGAA GAAATCACAC TGCCCTCAAG ACTCATATAT TACATCAACC AAGACTCGGA 600
AAGCCCTTAT CAGTCTCTTG ACACAAGGCG AAGACACCAG CAAAAACATA ATAAGGCTGT 660
CCATCTGCCC CAGGCAAGCT TCCAGATTGA AGCCTTCGGC TCCAAATTCA TTCTTGACCT 720
CATACTGAAC AATGGTTTGT TGTCTTCTGA TTATGTGGAG ATTCACTACG AAAATGGGAA 780
ACCAAGTATC TCTAAGGGTG GAGAGCAGTG TTAATACCAT GGAAGCATCA GAGGCGTCAA 840
AGACTCCAAG GTGGCTCTGT CAACCTGCAA TGGACTTCAT GGCACTTTTG AAGATGATAC 900
CTTGTGTAT ATGATAGAGC CACTAGAGCT GGTTCATGAT GAGAAAAACA CAGGTGAGCC 960
ACATATAATC CAGAAAACCT TGGCAGGACA GTATTCTAAG CAAATGAAGA ATCTCACTAT 1020
GGAAGAGAGT GACCATGGGC CCTTCTCTC TGAATTACAG TGGTTGAAA GAAGGAAGAG 1080
AGCAGTGAAT AACTCAGCTG GTATATTGTA AGAAATGAAA TATTTGGAAC TTATGATTGT 1140
TAATGATCAC CAAACGTATA AGAAGCATCG CTCTTCTCAT GCACATACCA ACAACTTTGC 1200
AAAGTCCGTG GTCAACCTTG TGGATTCTAT TTACAAGGAG CAGCTCAACA CCAGGGTTGT 1260
CCTGGTGGCT GTAGAGACCT GGACTGAGAA GSATCAGATT GACATCACA CCAACCTGT 1320
GCAGATGCTC CATGAGTTCT CAAAATACCG GCAGCGCAT TAAAGCATG CTGATGCTGT 1380
GCACCTCATC TCGCGGGTGA CATTTCACTA TAAGAGAAGC AGTCTGAGTT ACTTTGGAGG 1440
TGTCTGTCTT CGCACAGAGG GAGTTGGTGT GAATGAGTAT GGTCTTCCAA TGGCAGTGGC 1500
ACAAATATTA TCGCAGAGCC TGGCTCAAAA CCTTGGAAAT CAATGGGAAC CTCTAGCAG 1560
AAAGCCAAAA TGTGACTGCA CAGAATCCTG GGGTGGCTGC ATCATGGAGG AAACAGGGGT 1620
GTCCCATCTT CAAAAATTTT CAAAGTGCAG CATTTTGGAG TATAGAGACT TTTTACAGAG 1680
AGGAGGTGGA GCCTGCGCTT TCAACAGGCC AACAAAGCTA TTTGAGCCCA CGGAATGTGG 1740
AAATGGATAC GTGGAAGCTG GGGAGGAGTG TGATTGTGTT TTTCAATGGA AATGCTATGG 1800
ATTATGCTGT AAGAAATGTT CCTCTCCAA CGGGGCTCAC TGCAAGGAGC GGGCCTGTCT 1860
TAACAATACC TCTATGCTTT TTCAAGCAGC AGGGTATGAA TGGCGGATG CTGTGAACGA 1920
GTGTGATATT ACTGAATATT GTACTGGAGA CTCTGTGTCG TGCCACCAA ATCTTCATAA 1980
GCAAGACGGA TATGATGCA ATCAAAATCA GGGCGGCTGC TACAATGGCG AGTGCAAGAC 2040
CAGAGACAAC GATGTCTAGT ACATCTGGGG AACAAAGGCT GCAGGCTCTG ACAAGTCTGT 2100
CTATGAAAAG CTGAATACAG AAGGCACTGA GAAGGGAAGC TGGGGGAAGG ATGAGAGCCG 2160
GTGGATTGAG TGTGAGCAAC ATGATGTGTT CTGTGGATTC TTAATCTGTA CCAATCTTAC 2220
TCGAGCTCCA CGTATTGCTC AACTTCAGGG TGAGATCATT CCAACTTCTT TCTACCATCA 2280
AGCGCGGGTG ATTGACTGCA GTGGTGCCCA TGTAGTTTAA GATGATGATA CGGATGTGGG 2340
CTATGTAGAA GATGGAACGC CATGTGGCCC GTCTATGATG TGTTTAGATC GGAAGTGCCT 2400
ACAAATTCAA GCCTCAATTA TGAGCAGCTG TCCACTCGAT TCCAAAGGTA AAGTCTGTTC 2460
GGGCCATGGG GTGTGTAGTA ATGAAGCCAC CTGCATTGTT GATTTCACCT GGGCAGGGAC 2520
AGATTGCACT ATCGGGATC CAGTAGGAAA CCTTCACCCC CCAAGGATG AAGGACCCAA 2580
GGGTCTCTAG GCCACCAATC TCATAATAGG CTCATCTGCT GGTGCCATCC TGGTAGCAGC 2640
TATTGTCTCT GGGGGCAGAG GCTGGGGATT TAAAAATGTC AAGAAGAGAA GGTTCGATCC 2700
TACTCAGCAA GGGCCCATCT GAATCAGCTG CGCTGGATGG ACACCGCCTT GCACTGTGTT 2760
ATTCTGGGTA TGACATACTC GCAGCAGTGT TACTGGAACT ATTAAGTTTG TAAACAAAAC 2820
CTTTGGGTGG TAATGACTAC GGAGCTAAGG TTGGGGTGAC AAGGATGGGG TAAAGAAAAA 2880
CTGTCTCTTT TGGAAATAAT GTCAAAGAAC ACCTTTCACC ACCTGTCACT AAACGGGGGA 2940
GGGGGCAAAA GACCATGCTA TAAAAAGAAC TGTTCAGAAA TCTTTTTTTT TCCCTAATGG 3000
ACGAAGGAAC AACACACACA CAAAAATTAA ATGCAATAAA GGAATCATT AAAAA
```

Seq ID NO: 150 Protein sequence:  
Protein Accession #: NP\_003803

```
1      11      21      31      41      51
MKPPGSSSRQ PPLAGCSLAG ASCGPQRGPA GSVPASAPAR TPPCRLLLV LLLPPLAASS 60
RPRWGAAPAP SAPHWNETA EKNLGVLADE NTLQNNSSSN ISYSNAMQKE ITPLSRLIYY 120
INQDSESPVH VLDTKARHQK KHNKAVHLAQ ASFOIEAFGS KFIIDLILAN GLLSSDYVEI 180
HYENGKPYQS KGGECHYHSG SIRGVKDSKV ALSTCNGLHG MFDDETFVYM IEPLELVHDE 240
KSTGRPHIIQ KTLAQGYSKQ MNKLTMRGD QWPFLESLQW LKRRKRAVNP SRGIFEMKY 300
LELMIVNDHK TYKHKRSSHA HTNFAKSVV NLVDSIYKEQ LNTRVVLVAV ETWTEKDQID 360
ITTNFVQMLH EFKYRQRILQ QHADAVHLIS RVTFFHYKRSS LSYFGGVCSS TRGVGVNEYG 420
LPMVAQVLVS OSLAQNLGIG WEPSSRKPKC DCTESWGGCI MEETGVSHSR KFSKCSILEY 480
RDFLQRGSGA CLFNRPTKLF EPTCCNGYV EAGEECDCGF HVECYGLCKK KCSLSNGAHC 540
SDGPPCCNNTS CLFPQPRGYE RDAVNECDIT EYCTGDSGQC PPNLHKQDGY ACNQNGRCRY 600
NGECKTRDNQ CQYIWTAKAA GSKFKCYEKL NTEGTEKGNC GKDGDRWQC SKHDVFCGFL 660
```

WO 02/086443

LCNLTRAPR IGQLOGEIIIP TSFYHQGRVI DCSGARHVLD DTDVGVYVED GTPCGPSMMC 720  
LDRKCLQIQA LNMSSCPLDS KGVKCSGHGV CSNEATCICD FTWAGTDCSI RDPVRNLHPP 780  
KDEGPKGPSA TNLIGSIAI AILVAAIVLG GTGWGFKNVK KRRFDPQQG PI

PCT/US02/12476

Seq ID NO: 151 DNA sequence  
Nucleic Acid Accession #: NM\_023915  
Coding sequence: 250-1326

1 11 21 31 41 51  
GGCAGGAGG TTTCGTTTTC ATGCTTTACC AGAAAATCCA CTTCCTGACC GACCTTAGTT 60  
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120  
GTGAATGGAC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180  
CCACGSCCTC AATCGTCCCC AAGTGTTTCC TGACACGCAT CTTTGCTTAC AGTGCAATCAC 240  
15 AACTGAAGAA TGGGTTTCAA CTTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGGC 300  
CAGAGAGTC ACAATTTCAG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCTTCAC 360  
AATGAATTG ACACAATTGT CTTGCCGGTG CTTTATCTCA TTATATTGT GGCAGCATC 420  
TTGCTGAATG TTATTACAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA 480  
20 TTCTATCTCA AAAACATAGT GGTTCGAGAC CTCATAATGA CGCTGACATT TCCATTTTCA 540  
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600  
TCAGTTTTGT TTTATGCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT 660  
GATCGCTATC TGAAGGTGGT CAAGCCATTT GGGGACTCTC GGATGTACAG CATAACCTTC 720  
ACGAAGGTTT TATCTGTTTG TGTTTGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780  
25 ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAA ACTTAAAGT 840  
CCTTTGGGGG TCAATGGCA TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900  
GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960  
AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020  
GTGGCTGTGT TTTTACCTGT CTTTCTACCA TATCACTTGT GCAGAAATCC TTTTACTTTT 1080  
30 AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAGAA 1140  
ATTACACTTT TCTTGTCTGC GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTCATG 1200  
TGTAGGTGAT TTTCAAGAA GCTGTTCAA AAATCAAATA TCAGAACAG GAGTGAAAGC 1260  
ATCAGATCAC TGCAAGATGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTCACTGAT 1320  
GTGAGGCTC TTTATGTTT GTTGAATCG ATATGTACAA AGTGAAATA AATGTTTCTT 1380  
TTCATTATCC TTAATAAAAA AA

Seq ID NO: 152 Protein sequence:  
Protein Accession #: NP\_076404

1 11 21 31 41 51  
MGFNLTLAKL PNNELHGGES HNSGNRSDGP GRNTTLENEF DTIVLPVLYL IIFVASILLN 60  
GLAVWIFPHI RNKTSFIFYL KNIVVADLIM TLTFPPRIVH DAGFGPWYFK FILCRYTSVL 120  
FYANMYTISV FLGLISIDRY LKVVKEPQDS RMYSTITPKV LSVCVWVIMA VLSPNIIIT 180  
45 NGQFTEDNIH DCSKLKSPGL VKMHTAVTVV NSCLFVAVLV ILIGCYIAIS RYIHKSSRQF 240  
ISQSSRKRRK NQSIKRVVAV PFTCFPLPYHL CRIPFTPSHL DRLLDESAQK ILYYCKEITL 300  
FLSACNVCLD PIITYPFMCRS PSRRLEFKSN IRTSESIRS LQSVRRSEVR IYYDYDVT

Seq ID NO: 153 DNA sequence  
Nucleic Acid Accession #: D80008.1  
Coding sequence: 149-739

1 11 21 31 41 51  
GTTCCGCGCC AAAGCGCGGA GCGGAGGCGG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60  
CGAAAGGAGT GAGGCGCCGA GAGCCCGAGT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120  
AAGGCGCGCG GAGTGGGAAG CGTCCGCCAT GTTCTGCCGA AAAGCCATGG AACTGATCCG 180  
CGAGCTGCAT CGCGCGCCCG AAGGGCAACT GCCTGCCTTC AACGAGGATG GACTCAGACA 240  
60 AGTTCTGGAG GAGATGAAAG CTTTGTATGA ACAAACAGG TCTGATGTGA ATGAAGCAAA 300  
GTCAAGTGGG CGAAGTGATT TGATACCAAC TATCAAATTT CGACACTGTT CTCTGTTAAG 360  
AAATCGAGCC TGCACGTAGT CATACCTGTA TGACCGCTTG CTTGGATCA GAGCACTCAG 420  
ATGGGAATAT GGTAGCGTCT TGCCAAATGC ATTACGATT CACATGGCTG CTGAAGAAAT 480  
GGAGTGGTTT AATAATTATA AAAGATCTCT TGCTACTTAT ATGAGGTCAC TGGAGGAGA 540  
TGAAGGTTTG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGTCCG 600  
65 GTGTCTAATA GACTATGGAG AATTGGAAGT TGATGATGGC ACTTCAGTCC TATTAATAAA 660  
AAATAGCCAG CACTTTTAC CTGATGGAA ATGTAGCAG CTGATCAGAC AAGGAGTCCCT 720  
GGAGCACATC CTGTCTAGC CATGCGCGGA GGCACTTCCA GGCTTCACTC AACTCATGGA 780  
CTCCTCTGTA CTCACCTCTC CACCACTCC CTTCACTCC CTCTTTGATT TTGAAGTCA 840  
TAGACATTGT TTAAGATAAC TAAGAAATCT TGGCTAAGAA GTATAATTG CTAACATTA 900  
70 AGGACTTTCT TTTTAAATG TTGTACACTA TTCTTCTTAC TCTTTTGGG TTTTGGTTT 960  
GTTTTGTAGA GACTGTCTCA CTATGTTGCC CAAGCTGGCT TCAAACTCCT GGCTCAAGC 1020  
AGTCTTCCCA CTTTAGCTTC TCAAAGTGTG GAGATCAGAG GGTGAGCCA CTGCACCCGG 1080  
CCCTACTCC TTTTCTAAT AAGCTGTATC TGTAAATACA GCATTCTAC AGTTGTTACA 1140  
75 GTGTGTTTTT TAAATGAAAG TAAACATGTT TACATTGAA TCTCTTAAAT AAGCAGTCAC 1200  
TTGGCTGGAC AGGAAGAAGG TAGATCCTGT GTGTCTGTT TTCTGGTCAT GTGATTGTA 1260  
CAAGCTAGAG AGCTGAATTT CTGAGATACA CATTTCAAA TCACATGCAA GTGAAGATGA 1320  
TGGTCTGTAG AAATTTTCAG TATATATAAT GTTAAATGAC ATACTAATTT ATCATCTGGC 1380  
TATTTGGGAA GGAAGGACAC ACATGGATT TGCACATTTC CACCATGGTG GCTGGTGTGG 1440  
CTTGTGGCTA TGGGGTGATC ACCAGTATCA CCACCTTGGG AGGGACAGT GAAATGGGG 1500  
80 CTAGAGAAGG AACTTTGTAC AGTTTCCCT GAGATTGAGA TTGACTGAAA AGTCACATGA 1560  
AGAGTTGATT GTCTTTAAT GGTATGTTT AAACAGCTGA CATTTTAAAT TTTGATGAAA 1620  
TCCAGTTTAT TCGTTGTTC TTTTATGCTT TGGGTGTGTC ATCCGAGAAA TCTTTTCCCA 1680  
TCCCAAGATC ACAATTTTTT TTCCTTTTCA CTTCTAGAAG TGTATAAAT TTAAGCTTTA 1740  
85 TACTTTGGTC TATGACCCGT TTTTTTTTT GTTTTGTGTT GTTTTTCGTT TGTGTTCTTT 1800  
GTTTGTAGAT GGAGTCTGT TCTGTACCC AGGCTGGGGT GCAGTGGCGT GATCTGGCT 1860  
CACTGCAATC TCTATCCCT GGGTTCAAGT GATTCTCTG TCTCAGCCTC CCAAGTAGCT 1920  
GGGATTACAG GCACAGGCGG CCACGCTCGG CTAATTTTGT TATTTTATAG AGAGACAGAG 1980

TTTTACCATG TTGGCCAGGC TGGTTTCAAA CTCTGACCT CAAGTGACCC ACCTTGGCCT 2040  
 CCGCAAGTTT TGGGATTACA AGTGTGGGCC ACCGCGGCCA GCCTATGATC CATTTTGAAT 2100  
 GAATTTTTTA TATGTGCAAA GGTGTCAATC CACCTTCACT TTTTCTTGGG AATATAGATA 2160  
 TCCAGCTGTT TCACTACCAT TTTTGAAG GACTGCCCCT TGCTCTATCA CCTTGCATT 2220  
 TTTGTAAAA AGTAGTTGTC AATGTATATG TGGGTTTATT TCAGGACTCT GTTTTGTTC 2280  
 ATTGACCTGT TTTTCTCTCC TGAATGCCAA TACCATATTT GTATGTAGTG TATGTAAATT 2340  
 TCTAATAATT CTGAAACAG ATAGATTAA TGTGTATAT TTTTGTCTGT GTTTGTATT 2400  
 TTTGTAGAGA TGGGTTTCA CCGTGTGGC CAGGCTGTGT TGAACCTCTG AGCTAAAGCA 2460  
 ATACACTGTC CTGCTCTCC CCATGTGCTG GGATTACAGG CGTGAGCCTT GGTGCTGGCC 2520  
 CAGTGTACCA CATTTCTTTT TGAGATTGTT TTTGGCTATG TTAAGTCTTT TGCTTTTGAT 2580  
 GTGAAATTG GGAACAGGCA GGGTGTGGTG GCTTATGCCT GTAATCCTAG AACCTTGGGA 2640  
 GGCTAGATG GGTGGATCAC TTGAGCTCAG GAGTTCAGCA CCAGCCCGGG CCTATGGCAA 2700  
 AACTCCGTCT CTACAAAAA TAGAAAAAAT TAGCCAGGTG TGGTGGTGCA TGCCTGTAGT 2760  
 CACAGTTACA CGGCAGGCTG AGGTGGGAGG ATCACTTGAU CCCCAGAGGT CAAGACTGCA 2820  
 GTGAGCTGAG ATCACACCAC TGTACTCCAG CCTGGGTGAC AAAGTGAGAC TCTATCTCAA 2880  
 AAGAAATTA TGATCAATTT GTCAATTTCT ACAACAACAA CAACAAAAAC CCCTGTGGG 2940  
 CACCTTGATT GAGATTGAT TGAATTATA TAAACTGTT GGGAGAAATT ACATCTTAAT 3000  
 AATATTGAGT CTCTGCGCT ATAAACAAGG TCTGTCTCC TAGGTATTA TGTTTGTCT 3060  
 TCTATTCTC TTAATAATCT TTTGTAGTT TCAGTGTACA GGTCTACCAT GTCAGCATTT 3120  
 CATAGTTTG ATGCTAAATG GTATTTTAAA ATTTCAAATT CTAACCACCT GTTGCTAGTA 3180  
 AATAGAAATA CAATTGATG TGAACCTGTA TCCTTCAGCC TTGCTAACT GTGAGTTCTC 3240  
 ATGGTGTATT TGTAAATTAC ATCAACAGTC ATGTGTCTA TGAATAAAGA GTTTACTCC 3300  
 TTC

Seq ID NO: 154 Protein sequence:  
 Protein Accession #: BAA11503.1

1 11 21 31 41 51  
 MFCEKAMELI RELHRAPEQ LPAFNEGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60  
 TIKFRHCSLL RNRRCVAYL YDRLLRIRAL RWEYGSVLPN ALRPHMAEE MEFNNYKRS 120  
 LATYMRSLGG DEGLDITQDM KPKPSLYIEV RCLKDYGEFE VDDGTSVLLK KNSQHFLFRW 180  
 KCEQLIRQGV LEHLLS

Seq ID NO: 155 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 149-709

1 11 21 31 41 51  
 GTTCGGCGCC AAAGCGCGGA CGCGAGGCGG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60  
 CGAAAGGAGT GAGGCGCCGA GAGCCAGAT ACCATTTTGG CGTGAGAGCT GGTGGTTGGC 120  
 AAGGCGCGGG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCOG 180  
 CGAGCTGCAT CGCGCGCCCG AAGGCGCACT GCCTGCGCTT AAGGAGGATG GACTCAGACA 240  
 AGTTCTGGAG GAGATGAAG CTCTGTATGA ACACAAACAG TCTGATGTGA ATGAAGCAAA 300  
 GTGAGGTGGA CGAAGTGATT TGATACCAAC TATCAAAATT CGACACTGTT CTCTGTAAAG 360  
 AAATCGAGCG TGCACTGTAG CATACCTGTA TGACCGCTTG CTTCGGATCA GAGCACTCAG 420  
 ATGGGAATAT GGATGCGTCT TGCCAAATGC ATTAGCATTT CACATGGCTG CTGAAGAAAT 480  
 GGAGTGGTTT AATAATTATA AAGATCTCT TGCTACTTAT ATGAGGTGAC TGGGAGGAGA 540  
 TGAAGTTTGG GACATTACAC AGGATATGAA ACCACCAAAA AGCCTATATA TTGAAGCTGG 600  
 ATGAGTGGG GCGATCTCGG CTCAACCTGC AACCTCCACC TCCAGGTTTC ACCTCAACTG 660  
 CAACCTCCAC CTCCCAGGTC CGGTGTCTAA AAGACTATGG AGAATTGAA GTTGATGATG 720  
 GCACCTCAGT CCTATTAAAA AAAAATAGCC AGCACTTTT ACCTCGATGG AAATGTGAGC 780  
 AGCTGATCAG ACAAGGAGTC CTGGAGCACA TCCTGTCTAT ACCATCGGCC GAGGCACTTC 840  
 CAGGCTTAC TCAACTCATG GACTCCTCTG TACTCACTCT CTCCCACT CCCTTCACCT 900  
 CCCTCTTTGA TTTTAGAAGC TATAGACATT GTTAAAGATA ACTAAGAATA CTGGCTAAG 960  
 AAGTATAATT TGCTAACTAT TAAGGACTTT CTTTTATTAA TGTGTACAC TATTCTTCTC 1020  
 ACTCTTTTGT GGTTTTGGTT TTGTTTGTGA GAGACTGTCT CACTATGTTG CCCAAGCTGG 1080  
 TCTCAAACTC CTGGGCTCAA GCGATCTCTC CACCTTAGCT TCTCAAAGTG TTGAGATCAC 1140  
 AGGCGTGGAG CACTGCACCC GGCCCTACT CCTTTTCTA ATAAGCTGTA TCTGTAATCA 1200  
 CAGCATCTCT ACAGTTGTTA CAGTGTGTT TTTAAATGAA AGTAAACATG GTTACATTGG 1260  
 AATCTCTTAA ATAAGCAGTC ACTTGGCTGG ACAGGAAGAA GGTAGATCCT GTGTGTCTTG 1320  
 TTTCTGCTGC ATGTGTATTG TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA 1380  
 AATCACATGC AAGTGAAGAT GATGGTCTGT AGAAATTTTC AGTATATATA ATGTTTAATG 1440  
 ACATACTAAT TTATCATCTG GCTATTGGG AAGGAAGGAC ACACATGGAT TTGACACATT 1500  
 TCCACCATGG TGGCTGGTGT GGCTGTGGC TATGGGGTGA TCACCAAGTAT CACCACCTTG 1560  
 GAAGGGGACA GTGAAATTGG GGCTAGAGAA GGAACCTTGT ACAGTTTTC CTGAGATTCA 1620  
 GATTGACTGA AAGTGCACAT GAAGAGTTGA TTGTCTTTTA ATGGTATGTT TTAACAGCT 1680  
 GACATTTTAA ATTTGATGA AATCCAGTTT ATTCGTTTGT TCTTTATGTC TTTGGGTGTT 1740  
 GCATCCGAGA AATCTTTTCC CATCCCAAGA TCACAATTTT TTTTCTTTT TACTCTAGA 1800  
 AGTGTATATA TTTAAGCTT TATACTTGG TCTATGACCC GTTTTCTTT TGTTTTGTG 1860  
 TTGTTTCTTC GTTTGTTTCT TTGTTTGGAG ATGGAGTCTT GTTCTGTAC CCAGGCTGGG 1920  
 GTGAGTGGC GTGATCTGG CTCACTGCAA TCTCTATCCC CTGGGTCAA GTGATCTCT 1980  
 TGTCTCAGCC TCCCAAGTAG CTGGGATTAC AGGCACAGGC CCGCACGCTT GGCTAATTTT 2040  
 TGTATTTTTA GTAGAGACAG AGTTTACCA TGTGGCCAG GCTGGTTTCA AACTCCTGAC 2100  
 CTCAGTGAC CCACCTTGCC CTCCCAAGT TTTGGGATTA CAAGTGTGGG CCACCGCGGC 2160  
 CAGCCTATGA TCCATTGTA ATGAATTTT TATATGGTGC AAGGTGTCAA TCCACCTTCA 2220  
 CTCTTTCTTG GGAATATAGA TATCCAGCTG TTTCACTACC ATTTTGTGAA AGGACTGCCC 2280  
 TTTGCTCTAT CACCTTGTGA TTTTGTATA AAGTAGTTG TCAATGATA TGTGGTTTGA 2340  
 TTTGAGGACT CTGTTTGTG CCATTGACCT GTTTTCTCT CCTGAATGCC AATACCATAT 2400  
 TTGTATGTAG TGTATGTAAT TTCTTAATAA TTCTTGAAC AGATAGTATT AATGTGTCAT 2460  
 ATTTTGTGCT TGTTTGTAT TTTTGTAGA GATGGGGTTT CACCGTGTG GCCAGGCTGT 2520  
 GTTGAACCTC TGAGCTAAG CAATACACT GCCTCGTCT CCCATGTGC TGGGATTACA 2580  
 GCGGTGAGC TGGTGTGCG CCGAGGTGAC CACATTTCT TTTGAGATT GTTTTGGCTA 2640  
 TGTAAAGTCC TTTGCTTTTG ATGTGAAATT TGGGAACAGG CAGGGTGTGG TGGCTTATGC 2700  
 CTGTAATCCT AGAAGTTTGG GAGGCTTAGA TGGGTGATC ACTTGAAGCT AGGAGTTCCA 2760  
 GACCGCCCG GGCCTATGGC AAAACTCCGT CTCTACAAA AATAGAAAAA ATTAGCCAGG 2820

WO 02/086443

PCT/US02/12476

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

TGTTGGTGGT CATGCCCTGTA GTCCACAGTTA CACGGCAGGC TGAGGTGGGA GGATCACTTG 2880  
AACCCCGAGAG GTCAAGAGCTG CAGTGAGCTG AGATCACACC ACTGTACTCC AGCCTGGGTG 2940  
ACAAAGTGAG ACTCTATCTC AAAAAGAAAT TAGGATCAAT TTGTCAATTT CTACAACAAC 3000  
AACAAACAAA ACCCCTGTGG GGCACCTTGA TTGAGATTGC ATTGAATTTA TATAAACTG 3060  
TTGGGAGAAAT TGACATCTTA ATAATATTGA GTCTTCTGGC CTATAACAA GGCTGTCTT 3120  
CCTAGGTATT AATGTTTTGT CTTCTATTTC TCTTAATAAT CTTTTGTAGT TTTCAGTGTA 3180  
CAGGTCTACC ATGTGAGCAT TTCATAGTTT TGATGCTAAA TGGTATTTTA AAATTTCAA 3240  
TTCTAACAC TTGTGCTAG TAAATAGAAA TACAATTGAT GTTGAAGTTG TATCCTTCAG 3300  
CCTTGCTAAA CTGTGAGTTC TCATGGTGT TTTGTAAATT ACATCAACAG TCATGTGTTC 3360  
TATGAATAAA GAGTTTTACT CCTTC

Seq ID NO: 156 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
MPCEKAMELI RELHRAPEQG LPAPNEDGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60  
TIKFRHCSLL RNRCTTAVLY YDRLRLIRAL RNEYGSVLPN ALRPHMAAEE MEWFNNYKRS 120  
LATYMRSLGG DEGLDITQDM KPPKSLYIEA GCSGAISQAP ATSTSQVHLN CNLHLPGPVS 180  
KRLWRI

Seq ID NO: 157 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-621

1 11 21 31 41 51  
TTGGGGCCCA AAGCGCGGAG CGGAGSCCGA GCGGAGAGCC TGGCGCTGTA GGACTAGAAC 60  
GAAAGAGAGT AGCGCGCGAG AGCCAGAGATA CCATTTTGGC GTGAGAGCTG GTGGTTGGCA 120  
AGCGCGCGGG AGTGGGAAGC GTCCGCCATG TTCTGCGAAA AAGCCATGGA ACTGATCCGC 180  
GAGCTGCATC CGCGGCCCGA AGGGCACTG CCTGCCCTCA ACGAGGATGG ACTCAGACAA 240  
GTTCTGGAGG AGATGAAAGC TTTGTATGAA CAAAACCACT CTGATGTGAA TGAAGCAAAG 300  
TCAGGTGAGC AAGTGAATTT GATACCAACT ATCAAAATTC GACACTGTTC TCTGTAAAGA 360  
AATCGACGCT GCACGTGAGC ATACCTGTAT GACCGCTTGC TTGCGATCAG AGCACTCAGA 420  
TGGGAATATG GTAGCGTCTT GCCAATGCA TTACGATTTC ACATGGCTGC TGAAGAAGTC 480  
CGGTGTCTAA AAGACTATGG AGAATTTGAA GTTGATGATG GCACTTCAGT CCTATTAAAA 540  
AAAAATAGCC AGCAGCTTTT ACCTCGATGG AAATGTGAGC AGCTGATCAG ACAAGGAGTC 600  
CTGGAGCACA TCCGTGTCAT ACCATGCGCC GAGGCACTTC CAGGCTTCAC TCAACTCATG 660  
GACTCCTCTG TACTCACTCT CTCCACCACT CCTTCACCT CCTCTTTGA TTTTAGAAGC 720  
TATAGACATT GTTTAAGATA ACTAAGAATA CTGGCTAAG AAGTATAATT TGCTAACTAT 780  
TAAGGACTTT CTTTTTTTAA TGTGTACAC TATCTCTCT ACTCTTTTTT GGTTTTGGTT 840  
TTGTTTTGTA GAGACTGTCT CACTATGTTG CCCAAGCTGG TCTCAAATC CTGGCCTCAA 900  
GCAGTCTTCC CACTGTAGCT TCTCAAAGTG TTGAGATCAC AGGCGTGAGC CACTGCACCC 960  
GGCCCTACT CTTTTTTCTA ATAAGCTGTA TCTGTAATCA CAGCATTCTT ACAGTTGTTA 1020  
CAGTGTGTTT TTTAAATGAA AGTAAACATG GTTACATTTG AATCTCTTAA ATAAGCAGTC 1080  
ACTTGGCTGG ACAGGAAGAA GGTAGATCCT GTGTGCTTGG TTTTCTGTC ATGTGATTG 1140  
TACAAGCTAG AGAGCTGAAT TTCTGAGATA CACATTTTCA AATCACATGC AAGTGAAGAT 1200  
GATGGTCTGT AGAAATTTTC AGTATATATA ATGTTTAATG ACATACTAAT TTATCATCTG 1260  
GCTATTTGGG AAGGAAGGAC ACACATGGAT TTTGCACATT TCCACCATGG TGGCTGGTGT 1320  
GGCTGTGTCG TATGGGTTGA TCACCACTAT CACCACCTTG GAAGGGGACA GTGAAATTTG 1380  
GGCTAGAGAA GGAACCTTGT ACAGTTTTC CTGAGATTCA GATTGACTGA AAGTCACAT 1440  
GAAGAGTTGA TTGTCTTTTA ATGGTATGTT TTAACAGCT GACATTTTAA ATTTTGATGA 1500  
AATCCAGTTT ATTGCTTTGT TCTTTTATGC TTTGGGTGTT GCATCCGAGA AATCTTTTCC 1560  
CATCCCAAGA TCACAAATTT TTTTCTTTT TACTTCTAGA AGTGTATATA TTTTAAGCTT 1620  
TATACTTTGG TTTATGACCC GTTTTPTTIT TTGTTTGTIT TTGTTTTTTC GTTTGTCTT 1680  
TTGTTTTGAG ATGGAGTCTT GTTCGTGTC CCAGGCTGGG GTGCACTGGC GTGATCTTGG 1740  
CTCAGTCAAA TCTCTATCCC CTGGGTTCAA GTGATCTCT TGTCTCAGCC TCCCAAGTAG 1800  
CTGGGATTAC AGGCACAGGC CGCCAGCCTT GGCTAAATTT TGTATTTTAA GTAGAGACAG 1860  
AGTTTTTACA TTTTGGCCAG GCTGGTTTCA AACTCTGAC CTCAGTGAC CCACCTTGGC 1920  
CTCCCAAAGT TTGTGGGATTA CAAGTGTGGG CCACCGCGGC CAGCCTATGA TCCATTTTGA 1980  
ATGAATTTT TATATGGTGC AAGGTGTCAA TCCACCTTCA CTTTTTCTG GGAATATAGA 2040  
TATCCAGCTG TTTCACTACC ATTTTTTGAA AGGACTGCC TTTGCTCTAT CACCTTTGCA 2100  
TTTTTGTAA AAGTAGTGTG TCAATGTATA TGTGGGTTTA TTTCAAGACT CTGTTTTGTT 2160  
CCATTGACCT GTTTTCTCT CCTGAATGCC AATACCATAT TTGTATGTAG TGTATGTAAT 2220  
TTTCTAATAA TTCTTGAAC AGATAGTATT AATGTGTCAAT ATTTTGTCTG TTGTTGTAT 2280  
TTTTTGTAGA GATGGGTTT CACCGTGTG GCCAGGCTGT GTTGAACCTC TGAGCTAAAG 2340  
CAATACACTT GCTGTCTCT CCCCATGTC TGGGATTACA GCGGTGAGCC TTGGTGTCTG 2400  
CCCAGTGTAC CACATTTCTT TTTGAGATT GTTTTGGCTA TGTAAAGTCC TTTGCTTTG 2460  
ATGTGAATTT TGGGAACAGG CAGGGTGTGG TGGCTTATGC CTGTAATCCT AGAAGCTTGG 2520  
GAGGCTTAGA TGGGTGATC ACTTGAGCTC AGGAGTTCCA GACCAGCCCG GGCCTATGGC 2580  
AAAACTCCGT CTCTACAAA AATAGAAAAA ATTAGCCAGG TGTGTGGTG CATGCCTGTA 2640  
GTCAAGTTA CACGGCAGG TGAGGTGGGA GGATCACTTG AACCCAGAG GTCAAGACTG 2700  
CAGTGAGCTG AGATCACACC ACTGTACTCC AGCCTGGGTG ACAAGTGAG ACTCTATCTC 2760  
AAAAAGAAAT TAGGATCAAT TTGTCAATTT CTACAACAC AACACAAAA ACCCCTGTG 2820  
GGCACCCTGA TTGAGATTGC ATTGAATTTA TATAAACTG TTGGGAGAAAT TGACATCTTA 2880  
ATAATATTGA GTCTTCTGGC CTATAAACAA GGTCTGTCTT CCTAGGTATT AATGTTTTGT 2940  
CTCTATTTC TCTTAATAAT CTTTGTAGT TTTCACTGTA CAGGTCTACC ATGTGAGCAT 3000  
TTCATAGTTT TGATGCTAAA TGGTATTTTA AAATTTCAA TTCTAACAC TTGTGCTAG 3060  
TAAATAGAAA TACAATTGAT GTTGAACCTG TATCCTTCAG CCTTGCTAAA CTGTGAGTTC 3120  
TCATGGTGT TTTGTAAATT ACATCAACAG TCATGTGTTC TATGAATAAA GAGTTTTACT 3180  
CCTTC

Seq ID NO: 158 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51

MFCEKAMELI RSLHRAPEQ LPAFNEGLR QVLEEMKALY EQNQSDVNEA KSGGRSDLIP 60  
TIKFRHCSLL RNRCTVAYL YDRLLRIRAL RWEYGSVLPN ALRPHMAAEE VRCLKDYGEF 120  
EVDGTSVLL KNSQHFLEPR WKCEQLRQG VLEHILS

5 Seq ID NO: 159 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 149-229

10 1 11 21 31 41 51  
GTTGGCGGCC AAAGCGCGGA GCGAGGCGCG AGGCGAGAGC CTGGCGCTGT AGGACTAGAA 60  
CGAAAGGAGT GAGGCGCGGA GAGCCAGAT ACCATTITGG CGTGAGAGCT GGTGGTTGGC 120  
AAGGCGCGCG GAGTGGGAAG CGTCCGCCAT GTTCTGCGAA AAAGCCATGG AACTGATCCG 180  
CGAGCTGCTC CSCGCGCCCG AAGGGCAACT GCCTGCCTTC AACAAATTAGC TGGGTGTGGT 240  
15 GGCACACACC TGATGTCCCA GCAACTTAGG AGGCTGAAGT GAGAGGATTG CATGGCTCCA 300  
GGAAGTTGAA ACTGCACTGA ACTGTGGTCA CGCTATTACA CTCAGCCTG GGTGACAGAC 360  
TGAATCCCTG TCTCAAAAGG GAAAAGGAGG ATGGACTCAG ACAAGTTCTG GAGGAGATGA 420  
AAGCTTTGTA TGAACAAAC CAGTCTGATG TGTCTCTGT TAAGAAATCG ACGCTGCACT 480  
GTAGCATACC TGTATGACCG CTGCTTCGG ATCAGAGCAC TCAGATGG

20 Seq ID NO: 160 Protein sequence:  
Protein Accession #: Eos sequence

25 1 11 21 31 41 51  
ATGTTCTGCG AAAAAGCCAT GGAAGTATC CGCGAGCTGC ATGCGCGGCC CGAAGGGCAA 60  
CTGCTGCGCT TCAACAATTA G

30 Seq ID NO: 161 DNA sequence  
Nucleic Acid Accession #: U10694  
Coding sequence: 1333-2280

35 1 11 21 31 41 51  
GGATCCGGCC GGATCTCAGG GAGGTGAGGA CTTGTITCTC AGAGGGTGTG TGTGGACAAA 60  
ACAGGGAGGC CTTGTGTTGG ACAGACACAG TGGTCCCGAG ATTGGAGAGC AGTCCAGGTG 120  
AGGAACCTAA GGGAGGATCG AGGGTACCTC CAGGCCAGAG AAATCTCTAG ATCAAGAGAG 180  
TTTGCCTGCG CCTACTGTGC ACCCCAGAGA GCCCGGGCAG GCGTGTCTGC TGAGGTCCCT 240  
40 CCTTATCTCT GGGATCACTG GTGTGCGGGA GGGCTGGCCT TGGTCTGAGG GGGCTGCACT 300  
CACGTACAGA GAGGGAGGGT CCCAGGCCCT GCCAGGAGTC CAGGTGCGAGA CTGAGGGGAC 360  
CCCACTACCC AAACACAGAG GACCTAGCCC CACCCTGCCCT CTTGTGTGAG CTGAGGGGAG 420  
CGCTGGGGTG GATGACTTCC CCTCACTTCC TCTTCAGGTG TCTCTGGAG ATAGGGCCTC 480  
AGGTCAACAG AGGGAGGGTT CCAGACCTCG CAGGCATCAA GATGAGGACC AGGCAGTATC 540  
CTCACCCAGG GACACATGGA CCCCATTGAA TTAGACATC TCTTACTGTA CTTCCGAGGA 600  
45 AACCTGGGCG AGGTGTGGGC AGATGTTGGT TGGGGCATGT CCTTCTGTTC CATATCAGGG 660  
ATGTGAGCTC CTGATCTGAG AGACTCTCAG GCAAGTAGAG GAGTAGAGTC CAGTCCCTGC 720  
CAGGAGAAAG GTACGGGCCC TGAGTGAGCG CAGAGGGGAC CATCCACCCC AAAAGTGTGT 780  
AGAACTCAAG AGTGTCCAGC CGGCCCTCTT GACAGCACTG AGGGACCGGG GCTCTGCTG 840  
CAGTCTGAGC CCTAAGGGCC CCTCGATTCC TCTTCCAGGA GCTCCAGGAA GCAGGCAGGC 900  
50 CTTGGTCTGA GACAGTGTCC TCAGTTCGCA GAGCAGAGGA GACCCAGGCA GTGTCAGCAG 960  
TGAAGGTGAA GTGTTCAACC TGAATGTGCA CCAAGGGCCC CACCTGCCCC AGCACACATG 1020  
GGACCCCATG GCACCTGGCC CCATTCCCCC TACTGTCACT CATAGAGCCT TGATCTCTGC 1080  
AGGTAGCTCG CACGCTGAGT AGCCCTCTCA CTTCTCCCTC CAGGTTCTCG GGCAGGGCTA 1140  
ACGAGGAGGA CAGGAGCCCC AAGAGGGCCC AGAGCAGCAC TGAAGAGAC CTGTAAGTCA 1200  
55 GCCTTTGTTA GAACTCCAA GGTTCGGTTC TCAGCTGAAG TCTCTCACAC ACTCCCTCTC 1260  
TCCCAGGCCG TGTGGGTCTC CATCGCCCGC CTCTGCCCA CGCTCCTGAC TGCTGCCCTG 1320  
ACCAAGTCA TCACTGTCTC CGAGCAGAGG AGTCCGCACT GCAGGCTGTA TGAAGACCTT 1380  
GAAGCCCAAG GAGAGGACTT GGGCTGTATG GGTGCACAGG AACCCACAGG CGAGGAGGAG 1440  
GAGACTACCT CCTCTCTGAG CAGCAAGGAG GAGGAGGTGT CTGCTGCTGG GTCATCAAGT 1500  
60 CTTCCCCAGA GTCTCAGGG AGGGCTTCC TCCCTCATTT CGTCTACTA CACTTTATGG 1560  
AGCCAAATCG ATGAGGGCTC CAGCAGTCAA GAAGAGGAAG AGCCAAAGCTC CTCGGTCCGAC 1620  
CCAGCTCAGC TGGAGTTCAT GTTCCAAGAA GCACTGAAAT TGAAGGTGGC TGAGTTGGTT 1680  
CATTTCTGCG TCCACAAATA TCGAGTCAAG GAGCCGGTCA CAAAGGCAGA AATGCTGGAG 1740  
AGGTCTATCA AAAATTACAA GCGCTACTTT CCTGTGATCT TGGGCAAGAC CTCGAGTTTC 1800  
65 ATGCAAGTGA TCTTGGCAC TGATGTGAAG GAGGTGGACC CGCGCGGCCA CTCCTACATC 1860  
CTTGTCTGCT CTTTGGCCT CTGCTGCGAT AGCATGCTGG GTGATGGTCA TAGCATGCCC 1920  
AAGGCCGCGC TCTGATCAT TGTCTGGGT GTGATCCTAA CCAAGACAA CTGCGCCCTC 1980  
GAAGAGGTTA TCTGGGAAGC GTTGAAGTGT ATGGGGGTGT ATGTTGGGAA GGAGCACATG 2040  
70 TTCTACGGGG AGCCCGAGGA GCTGCTCACC CAAGATTGGG TGCAGGAAAA CTACCTGGAG 2100  
TACCGGCAGG TGCCCGCGAG TGATCCTGCG CACTACGAGT TCCTGTGGGG TTCCAAGGCC 2160  
CACGCTGAAA CCAGCTATGA GAAGGTGATA AATTATTTGG TCACTGCTCAA TGCAAGAGAG 2220  
CCCATCTGCT ACCCATCCCT TTATGAAGAG GTTTTGGGAG AGGAGCAAGA GGGAGTCTGA 2280  
GCACCGAGCG CAGCGGGGGC CAAAGTTTGT GGGTCAAGG CCCCATCCAG CAGCTGCCCT 2340  
75 GCCCATGTG ACATGAGGCC CATCTCTCGC TCTGTGTTTG AAGAGAGCAA TCAGTGTCT 2400  
CAGTGGCAGT GGGTGAAGT GAGCACACTG TATGTATCT CTGGTTCCT TGTCTATTGG 2460  
GTGATTTGGA GATTTATCTT TGCTCCCTTT TGGAAATGTT CAAATGTTCT TTTAATGGTC 2520  
AGTTTAATGA ACTTCAACAT CGAAGTTAAT GAATGACAGT AGTCACACAT ATTGCTGTTT 2580  
ATGTTATTTA GGAGTAAGAT TCTTGTCTTT GAGTCACATG GGGAAATCCC TGTATTTTTG 2640  
80 TGAATTGGGA CAAGATAACA TAGCAGAGGA ATTAATAATT TTTTGAAC TTGAACCTAG 2700  
CAGCAAAATA GAGCTCATAA AGAAATAGTG AAATGAAAT GTAGTTAATT CTGCTTAT 2760  
ACCTCTTTCT CTCTCTGTA AAATTAATAA ATATACATGT ATACCTGGAT TTGCTTGGCT 2820  
TCTTTGAGCA TGTAAAGAA ATAAAAATTG AAGAATAAT TTTTCTGTT CACTGGCTCA 2880  
TTTTTCTCTC AGACAGCAC TGAACATCTG TTATTCGGA CACCTGGGT T

85 Seq ID NO: 162 Protein sequence:  
Protein Accession #: AAA68877.1

1	11	21	31	41	51	
MSLEQRSPHC	KPDEDLBAQG	EDLGLMGAQE	PTGEEBETTS	SSDSKKEEVS	AAGSSSPFQS	60
PQGGASSSIS	VYITLWQFDP	EGSSSQEERE	PSSSVDPAQL	EFMPQELRL	KVAVELVHLL	120
HKYRVKEPVT	KAEMLSEVIV	NYKRYFPVIF	GKASEFMQVI	FGTDVKEVDP	AGHSYILVTA	180
LGLSCDSMLG	DGHSMPIKAL	LIIVLGVILT	KDNCAPEEVI	WEALSVMGVY	VGKEHMFYGE	240
PRKLLTQDWV	QENYLEYRQV	PGSDPAHYEP	LWGSKAHAET	SYEKVINYL	MLNAREPICY	300
PSLYEEVLGE	EQEGV					

Seq ID NO: 163 DNA sequence  
Nucleic Acid Accession #: AF292100  
Coding sequence: 30-809

1	11	21	31	41	51	
GGGGGGGGAG	AGGCCTGGAG	GACACCAACA	TGAACAAGTT	GAAATCATCG	CAGAAGGATA	60
AAGTTCGTCA	GTTTATGATC	TTCACACAAT	CTAGTGAAAA	AACAGCAGTA	AGTGTCTTTF	120
CTCAAAATGA	CTGGAGGTTA	GATGTTGCAA	CAGATAATTT	TTTCCAAAT	CCTGAACTTT	180
ATATACGAGA	GAGTGTAAAA	GGATCATTTG	ACAGGAAGAA	GTTAGAACAG	CTGTACAATA	240
GATACAAAGA	CCCTCAAGAT	GAGAATAAAA	TGGAATAGA	TGGCATAACG	CAGTTCCTGT	300
ATGACCTGGC	ACTCGATCCA	GCCAGCATT	GTGTGTTGAT	TATTGCGTGG	AAGTTCAGAG	360
CAGCACACA	GTGCGAGTTC	TCCAAACAGG	AGTTCATGGA	TGGCATGACA	GAATTAGGAT	420
GTGACAGCAT	AGAACCACTA	AAGGCCCAGA	TACCCAGAT	GGAACAAGAA	TTGAAAGAAC	480
CAGGACGATT	TAAGGATTTT	TACCACTTTA	CTTTTAATTT	TGCAAGAAT	CCAGGACAAA	540
AAGGATTAGA	CTCTAGAAATG	GCCATTGCCT	ACTGGAACCT	AGTGCCTAAT	GGAAGATTTA	600
AATCTCTAGA	CTTATGGAAT	AAATTTTGT	TGGAACATCA	TAAACGATCA	ATACCAAAAG	660
ACACTTGGAA	TCTTCTTTTA	GACTTCAGTA	CGATGATTGC	AGATGACATG	TCTAATTATG	720
ATGAAGAAGG	AGCATGCGCT	GTTCTTATTG	ATGACTTTGT	GGAATTTGCA	CGCCCTCAAA	780
TTGCTGGGAC	AAAAAGTACA	ACAGTGTAGC	ACTAAAGGAA	CCCTTTAGAA	TGTACATAGT	840
CTGTACATA	AATACAACAG	AAAATTGCAC	AGTCAATTTT	TGCTGGCTGG	ACTGAACATG	900
AGATCAATCC	TCACAATTTA	GACTGAGGGT	TGAGACAAAA	CTTTAAGGAT	ACATCTTGGA	960
CCATATCGTA	TTTCACTTCT	CTAATGGTGG	TTTGGGCTTG	TCTTCTAGTC	TGGGCGCTC	1020
TAAACATTTA	TAATTTCCAAC	ATTGTGGATT	TCATCTTATA	TCTGTGGACC	ATCCTAGTTF	1080
ATTCTCCCAT	AAGTCTTAGA	AGCTTTATGG	TGATTATTTT	GAGGTTTTC	TTCTCGCATA	1140
AAGCACAAATG	CTGTCTTCAT	CAGAAAACAG	TGGCATAAG	AATTAACAT	ATGAACATCA	1200
CAAAACAATT	TATAAAACT	TCTTAAATAT	ACGCTTTGGG	CTAGTTGCAA	AGACTATGCT	1260
AATAGCACTT	CCAGTGAAG	TGATATATTT	AAGTGTACTG	GATCTGGAAT	GGTGTTTTGG	1320
TTTGGGGGGA	ATTTTTTTTT	TTTCTGGCA	AATCACAAT	GTGTTGATG	TGAGTATCTG	1380
ATGAAAAAAC	AATGTCAGAA	TAAACGACAT	GAAAAATTTT	TAGGATAACT	TGGTGCCTAC	1440
CTGAAAAAATG	TATGTGTTT	TAGACTCTTG	ATTTCAAAG	GTTCACAGA	ACTAGTCTCG	1500
GCCTACCTTA	CCCATGTTTA	TATATAGCTG	TCCTACAGGG	AGCTTTTAT	TAGAAAATGT	1560
CTGCATAATG	TGATGTTCTT	CTCCTGTCTA	CATTATGCAC	TACATAATTG	GACTTCATTA	1620
TGCTTTTGAA	ATGCTTATCT	GCTGTGCACA	TAAGTTAAAC	TATTTAATTT	GTTTTGAATG	1680
TTTTGGATTG	CTACACAATA	CAATATTCTA	AATTTAGGCA	TGAGGGTTTT	TTTGTTTTAT	1740
TTTTTACTTT	TTTTTGTCT	TGCACATGG	AACACAAATG	AAATTTCTCT	AATTTATAAG	1800
AAGTAGTAG	GAGTTAAAT	TGAAAAATGG	TGTGATGAG	CCACGAAAT	CAATCTTTAT	1860
AATATAGGTA	CTGCTCTTTC	AGACAAACAG	TCCATTTT	ATGACTTCTT	ATTTTGTGTA	1920
AATTACTTTA	ACTGCTAATC	ACTGTGGTGG	CCAAATATTT	ACTTCAGAAG	CAAGATTTT	1980
CAAAACAAGCA	TACAGATGTC	AAAATACCAG	TCTGGCTTCT	AGTCTATTTA	CTGTTTGTGT	2040
TCACTCAGAT	TAGCTCAGTT	TCTCATCAA	AGCAGAAATG	TATCTTGGT	GTGTGTGTGT	2100
GTGTGTGTGT	GTGTGTGTGT	GTATGTGTGT	ATATATATAT	ATATATATAT	ATATATATTT	2160
TTTTTTTTTT	TTTTTTTTTA	ATTACAAAAG	CCATGAGCTG	CTTTTATGCT	GAAAATGGTC	2220
ATTTCCCTGT	TCACTTACTG	ACATGTGAAG	AAGGGTTTCT	TGCTTTCTTA	AACATTTCCG	2280
TAAGGCAGGC	TAGAAATGTA	ATACCTCAAA	TGTTTATGTA	TTATGGTCTT	TTGATAGGAA	2340
TAGATTCTGC	TGGGATATA	TATCCAGGCA	CTCTTAAGG	TCTAGGGTGT	ATATTAAACA	2400
AGGAATGTAC	TAGAAATAGC	AGTACATTTT	ATGCAATAT	GGAATTTAT	TTAAGAAACA	2460
ATGACATATC	ATAACTGCTT	TTTACATGAT	TTTGAATAG	ACTAGAAAGC	TTTCCCTATA	2520
GACATATATA	TATTTCCAATC	ATAACTTTAA	TTCAAGAAAT	CAGTTTATCC	AAAAGAAAAA	2580
TTTGAATTT	TCTATTGAGG	CTACTGGAAT	TGGTTATTAA	AAGAAAAAGG	AAAAAGAAAG	2640
ATCTTGCTGC	TTTCAGTATT	TCCTGATTTT	TTTGTAAATA	TAAAGAGGAA	CTTCAATTAT	2700
GAAAAATTTT	TAAAGATAT	ATATATCTAT	ATATCTATAT	ATATGTACTG	TTTTGTTCCT	2760
TGTCTTGAAG	ATTTTGAATT	ATGTTTATG	GTTCAGATT	GATTAATTCA	CATATGCTGT	2820
GTTTCTTTTA	AAAGTCATAT	GGGTTCTGG	CCTAATGCCT	TGGAATTTAC	ATATTTTCT	2880
TTTTAAATGC	AAAACCTTTT	CAACAAAATA	GTGTTTGTCA	TCAGGTTGGT	ACTAAACATT	2940
TATAATTACT	GTGTAATTAT	AAACAAAAAT	ACATAAAGCT	TTGAATATAA	TTATGTAGCA	3000
TAAAAGTTAA	GGTGTTCAC	TATGATGGCA	TCTTAGAAT	AAACAAAAC	TTTACTAGGG	3060
CTGAAAAGAG	AAGACTGATT	TAATGTGGTG	TGATTATTCT	GAAGATAAAT	GTCTGGCTAC	3120
AGGGAAATAT	TTGTAATAA	AAATGATTAC	ACATATGGCT	GTGTGTGTTT	GAGTCTGTGT	3180
CTGTGAGAGA	GCCAGAGAGA	GTGAGAGAGA	ITGACAGAGA	AAGGGAGAGA	CACACACACG	3240
CCCTTTGAAT	TGCTTTAACT	CCTAAGTGT	TCAGTCTCA	TTCCGGTAAA	CTCCCATGCT	3300
TGATTCTTTG	TTTTAAACTG	AACCATAGGT	ACAGTTTCT	TTTTGCCAAA	TGTCAAAACA	3360
GGTACAAAT	TTAAATGTG	ATGCTTTT	AATAGAAAAA	TGTATAAAAT	TAGAAAGTGC	3420
CACATATAAA	AAATACTTGA	GATGAAGATT	ATCTTTAGTG	AATATCATCT	GCATATCTCT	3480
GTAAGTTCAA	TTGCTTTCT	TACAGTCCCT	GTCTATTATC	CAACAGAGGC	AATAAAGCT	3540
GCAGTGAAT	TG					

Seq ID NO: 164 Protein sequence:  
Protein Accession #: AAG00606

1	11	21	31	41	51	
MNKLKSSQKD	KVRQFMFTQ	SSEKTAVSCL	SNQDWKLDVA	TDNFFQNPFL	YIRESVKGSL	60
DRKLEQLYN	RYKDPQDENK	IGIDGIQFQC	DDLALDPASI	SVLIIAWKFR	AATQCEFSKQ	120
EFMDQTELG	CDSEIQLKQ	IPKMEQELKE	PGRFKDFYQF	TNFKNPQPG	KGLDLEMAIA	180
YNNVLNLRP	KFLDLNKNFL	LEHHRSPK	DTWNLLDIFS	TMIADMSNY	DEEGAWFVLI	240
DDFVEFARPQ	IAGTKSTTV					

Seq ID NO: 165 DNA sequence  
Nucleic Acid Accession #: AF256215  
Coding sequence: 220-2028

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1	11	21	31	41	51	
CTCCAGTCCG	CATGCTCAGT	AGCTGCTGCC	GGCCGGGCTG	GGGGGGGGGG	TCCGCTGCGC	60
GCCTACGGGC	TGCGGTGGCG	GCCGCCGGCG	CACCCGGCAG	GGCCGGCCAG	TCCCGGCTTC	120
CCTGCTCCAG	AGCCGCCGCC	TGGGCCGGGG	CAGGGCGGGC	CCGGGGCTCC	TCCATGCTGC	180
CAGCCGCCGG	GCTGCGGAGC	CGACCAAGTG	GCTCCTGCGA	TGGCGGCGGA	AGAGGAGGCT	240
GCGCGGGGAG	GTAAGTGT	GAGAGAGGAG	AACCACTGCA	TGCTCTGT	GGTTCCAGC	300
CGCGTGATC	CAGGACAAG	ACCAACAGCT	ATGGGGTCTT	TCAGCTCACA	CATGACAGAG	360
TTTCCACGAA	AACGCAAGG	AAGTGATTCA	GACCCATCCC	AAGTGAAGA	TGGTGAACAC	420
CAAGTTAAAA	TGAAGGCCTT	CAGAGAAGCT	CATAGCCAAA	CTGAAAAGCG	GAGGAGAGAT	480
AAAAATGAATA	ACCTGATTGA	AGAACTGTCT	GCAATGATCC	CTCAGTGCAA	CCCCATGGCG	540
CGTAACTCGG	ACAACTTACG	AGTTTAAAGA	ATGGCTGTTC	AACACTTGAG	ATCTTTAAAA	600
GGCTTGACAA	ATTCCTATGT	GGGAAGTAAT	TATAGCCAT	CATTCTCTCA	GGATAATGAG	660
CTCAGACATT	TAATCCTTAA	GACTGCAGAA	GGCTTCTTAT	TTGTGGTTGG	ATGTGAAGCA	720
GGAAAAATTC	TCTTCGTTTC	TAAGTCAGTC	TCCAAAATAC	TTAATTATGA	TCAGGCTAGT	780
TTGATGAGAC	AAAGCTTATT	TGACTTCTTA	CATCCAAAAG	ATGTTGCCAA	AGTAAAGGAA	840
CAACTTCTCT	CTTTTGATAT	TTACCAAGAA	GAAAAGCTAA	TAGATGCCAA	AACCTGGTTG	900
CAAGTTCACA	TAATCTCCA	CGCTGGAAGG	ACACGTGTGT	ATTCTGGCTC	AAGACGATCT	960
TTTTTCTGTC	GGATAAAGAG	TTGTAAAATC	TCTGTCAAAG	AAGAGCATGG	ATGCTTACCC	1020
AACTCAAAGA	AGAAAGAGCA	CAGAAAATTC	TATACTATCC	ATTGCACTGG	TTACTTGAGA	1080
AGCTGGCCTC	CAAAATATTGT	TGGAATGGAA	GAAGAAAGGA	ACAGTAAGAA	AGACAACAGT	1140
AAITTTACCT	GCCTTGTGGC	CAITGGAAGA	TTACAGCCAT	ATATTGTTC	ACAGAACAGT	1200
GGAGAGATTA	ATGTGAACCC	AACCTGAATT	ATAACCCGCT	TTGCACTGAA	TGGAAAATTT	1260
GTCTATGTAG	ATCAAAGGGC	AACAGCGATT	TTAGGATATC	TGCTCAGGA	ACTTTTGGGA	1320
ACTTCTTGT	ATGAATATT	TCATCAAGAT	GACCACAATA	ATTGACTGA	CAAGCACAAA	1380
GCAGTTCTAC	AGAGTAAGGA	GAAAATACTT	ACAGATTCTC	ACAAATTCAG	AGCAAAAGAT	1440
GGCTCTTTTG	TAACCTTAAA	AAGCCAATGG	TTTAGTTTCA	CAATCCTTG	GACAAAAGAA	1500
CTGGAATATA	TGTATCTCTG	CAACACTTTA	GTTTGGGAC	ATAGTGAGCC	TGGAGAAGCA	1560
TCATTTTTAC	CTGTAGCTC	TCAATCATCA	GAAGAATCCT	CTAGACAGTC	CTGTATGAGT	1620
GTACCTGGAA	TGCTACTGG	AACAGTACTT	GCTGTGGTGA	GTATTGGAAC	AGATATTGCA	1680
AATGAAATTC	TGGATTTACA	GAGGTTACAG	TCTTCTTCAT	ACCTTGATGA	TTGAGTCCA	1740
ACAGGTTTAA	TGAAGATGAT	TCATCTGTGA	AACCTGCAGGA	GTATGTCAA	TAAGGAGTTG	1800
TTTCCACCAA	TGCTCTCTGA	AATGGGGGAG	CTAGAGGCTA	CCAGGCAAAA	CCAGAGTACT	1860
GTGCTGTCC	ACAGCCATGA	GCCACTCTCT	AGTGATGGTG	CACAGTTGGA	TTTCGATGCC	1920
CTATGTGACA	ATGATGACAC	AGCCATGGCT	GCATTATGTA	ATTACTTAGA	AGCAGAGGGG	1980
GGCTGGGAG	ACCTGGGGA	CTTCAGTGAC	ATCCAGTGGG	CCCTCTAGCC	TTTGATTTTT	2040
AACTCACAAA	ATGAGAAACA	TTTTAAAGCA	TTATTTACGA	AAAACTGTIC	TCACTATTTC	2100
TAAAGTACTG	TATTGATATT	GTTTGTATCT	TTTATTAATG	TTCTACCACT	TTTTATAGAT	2160
TTGCACTTTC	CTGTACAGG	GATGTGGGGA	AATACGTTTT	CTCCTCAAGA	GAACCAAGTT	2220
TATTATAGAC	TCCTTTATTC	AGTGAATGG	CTTATAATCC	ACTAGTTGCC	ATATTTTTCG	2280
TAAATATTT	CTAACCAAGA	ATACTACTTA	CATATTGTTT	TGGCTTTGTT	TTATTTTTCG	2340
TGCAGTTTTT	TTTGTGTGAG	GTAATGTAAT	ATATTGATGT	TTTCCTTTGT	GTCTAAGATT	2400
GATTATTAAT	AGTAGGTTTG	TATAATTGG	AACATTTTCC	ATGCCCTTGC	AAITTTCTTTA	2460
ATTGAGGATA	GGGCTTACAC	ACTTTAAGAA	AACAGTGAGT	ACTTGAACAT	TTAAGGGGAC	2520
AGTGCAATTT	ATAGTCATAA	TCACATTGAA	TACTGTATTT	GATCTTTGGA	GACTTAGGCA	2580
AGCACAGAGC	TGGATATTTT	ATGCTCAGTT	GAGCACTTTA	AGATGAATTT	TAAGTGAGAT	2640
GATTTCTTGC	TTAAACTCA	GAAGTCAAA	AGAGTTTCAG	CTTTCCTTAC	AGAAAAGGAA	2700
GGATCTTGGG	CCCTAGATCT	TGGGAGTTAA	CCTCTGCATA	TAAGATTTAC	TCTTAAATAGG	2760
CCAGAGCTGG	TGCTACGGCC	TGTAATCCCA	GTACTTTGGG	AGGCTGAGAC	GGGAGATCA	2820
CTTGAGGTCA	AGAGTTCAAG	ACCAGCCTGG	CCAATATGGT	GAACCCCGGT	TTCTACTAAA	2880
AATCAAAAAA	AAATTAACCA	GGCACTCACT	CTTGAGGTAA	CTAACCAACT	CCACGATATA	2940
TGACAGTCCA	TTCATGAGCG	CAAGGCCTC	ATGACCTAAT	GGCACACACC	TGTAATCCCA	3000
ACTGCTTGGG	AGGCTGAGGC	GAGAGGATTG	CTTGAACCTG	GGAGGCAGAG	GTTGCACTGA	3060
GCGGAGATCG	CACCACTGCA	CTCCAGTCTG	GGCAACAGAG	TGAGACTTCA	TCTCAAAAAA	3120
AGTAAAAAAA	AAGATTAAAT	ATAATCACTG	AAGATCTCTA	TTATAGATAG	AITAGGTTTT	3180
TGACATTGGA	AACATACTTA	GGGATAGATT	TGTCCTAAAG	GAATAAAGTA	GGCCCGGGCA	3240
GATTAAATGT	CTTGTGTAAA	GTCAACATT	AAATTCAGTC	ACACATTAAA	TTCATAGAGT	3300
TTTAAATGTT	TAATGTATAT	AAACCAGTTT	CTTTATACAC	ATTTGGGAAA	ACATTGGTCT	3360
CACAGATTAA	ATGATTAACT	AACAGCCCA	GGAAGTATGT	GTAGCTTTCT	AAGTAAATAG	3420
GCAATTACAG	TTATTGCCCT	TAACCAAGG	TAATAAAACA	AAATGACAA	TACATGTTTA	3480
AAATATGAG	GCAATGAGAA	ATAATTAA	AAACCAATTT	CTAGTTATAA	TTTAAATTTT	3540
GGAGAGCATT	TTTAAAGTA	ATTAATCCAG	AGGTGGCTCA	AATTGAGTAT	AAGAATTAAG	3600
ATTATTAA	ATACTGCATG	TCTACCTTCT	CGGGATCAT	ACTTTATAAC	ACTTCTCTCT	3660
TCAGTAGCTC	TTATAGCTTT	GCCAAAGTAT	CTCCCATATT	TTCTCTCTCG	TGCTCTGCAA	3720
ATGAAAGTCA	GATAGGCTGG	GAACCTATGG	GGCAGCCCTC	AGACTTCAAT	GTGGGCTTCA	3780
AATCCAGTTT	CCTGTTCTAT	ATGGTGCTAC	ATCITTCCAG	AAAAITTCCT	TCAGAGCCCC	3840
TCGCCAAAAA	AAAGCATTAT	TTTGACCTCG	CATGCTATTT	CTTAGCTGT	AGGTGATAGA	3900
TTAGAACTTC	TGTCAGACAT	GTTAATGACA	AACATACCAA	CAGACAATAA	CCAAAGCAAA	3960
TTTTCCTTTC	AAGTGTGAAA	TGTGCAGGGG	CTCGTGGGCA	AGGATGTATT	GGCACACTGT	4020
CCTCTTGAAC	TGATAGTGTG	CCAGCAATGT	TGGAGGTTGG	CACCAITTCCT	GGTCCGACAC	4080
TTGAGGACCT	GAGAGACATC	AGGTTTAGAA	TGAGCCAAAG	AAATCTTACA	AGATGGGGAG	4140
AATGTGTGTG	CAGCAGCCTA	AGTGTATAG	TTAAGTCTAA	AGAAGTATGA	AAGATCCCCCT	4200
GTGTCTCTTA	AAATTGAGCAG	AGGGGCTGCT	CTACCAATAT	CACITTTTAG	GGGACTGAAC	4260
CATTGCAAGT	TAGACTTGGC	TTCCAAAGAG	TCTGCCTAAG	CCAGGGGTGG	CAGGGTAGGC	4320
CATCATAGCT	GGATGCGCTC	AAAAGCAGAT	GGGGGCAGAC	TTGCCCTCGT	GATGCCAGGA	4380
TTTGAGAGGC	AGAGTTTCTA	GAGGGAGACC	AGTGCTGCCT	CTCACAGTGG	CAGTTTTTTC	4440
TCTTTGCAAG	AGGAGGGGCT	GTTCAATTCC	ATAGACCAAGT	GGGCAGATAG	CCAGTTGAAT	4500
ACTCTGTGCA	TGGTTTGTAT	CTTTATTAGT	TCGCTCTAAT	ATTTTCTGT	AGATCCTTTT	4560
GTCTCTGACT	CAAAATCTAA	TCCATGCATT	GTATGATACC	GTAGCTCTCC	TAAGGTTTGT	4620
GTTTCCTTCA	AAATGTTTTA	GTTTCTTCTA	ACTAAATTTG	ATTTTGTCTG	TTAGAAGTGA	4680
CATATTTTTA	TGGTATACAC	TATGTTCTCT	TTTTCTACTG	CGAGTCAATT	TTTTGAATTT	4740
TCGTGAGAAA	GAATATATCT	ACAAATTGCA	CGAAAGTATC	ATAAAACAGC	TACTCTAGAG	4800



CAGCGCTGTC CAATAGAAAT ATAATCTGAG CCACATGTAT AATTTTATTT TCTTCTAGCC 4860  
ACATTAAAGA AGTAAAAAGA TACAAGTAGA ACTAATTTTA ATGTTTATAT TCAGTATATC 4920  
CAAAATATCA TTGAAACATG TAATTAATAT AAAATTATTA ATGTGATATT TTACATTCTT 4980  
TTGGTAATAC TAGTCTTCAA AATCTGGTAT GTATCTTACA TTGATAGCAC ATCTCACTTT 5040  
GTACTAGCCA CATTGCAAGT GCTCAGTAGC CACATGTGGC TAGTGGCTAC TGCACCTGGAC 5100  
AGCAGATTC TAGGTTCCAC CCTAACACCC AAGTCTCTGT GATTAGAAAT CCAGAAATCAG 5160  
AGCTGGAAGT AACATAGAG ATCAAACTTC CTTTTAAAAA TGAGGACGCT GAGGCACAGA 5220  
GTTTAAATGG CTTGCAATGAG GTACATACAGC TAAATTCAGC CTCAACAGGG TCTTCTGATT 5280  
CCAGGCATCT TTCCCACTCC ACTACATTAC TGTAGTGGTA ATTCTTAGGG TTAATAAAAG 5340  
TGTAGAGTAG GCGCGGCGCA GTGGCTCATG CCTGTAATCC CAGCACTTTG GGAGGCCGAA 5400  
TGCGGCGGAT CAGCAGGTCA GGAGATCGAG ACCATCCTGG CCAACATGGT GAAACCCCGT 5460  
CTCTACTGAA AATACAAAGC AAAATTAGCC AGGTGTGGTG GCGGCGCCT GTGTCCTCAG 5520  
CTGCTCTGGA GGCTGAGGCA GAATGGCGTG AACCCAGGAG GCAGAGATGG CAGTGAGCCA 5580  
AGATGCGGCC ACTGCAACCC AGCCTGGGCG ACAGAGCGAG ACTCCATCTC AAAAAAAG 5640  
AAAAAAGAAA AAGAAAGAAA AAGAAAGTC TAGAGAATAT TATATTAGT GGTATTATT 5700  
GAAGTAGACC AATGTTTACC CCATAAGGAT ATTTTCTCTT AAATACCATG TTGGAAGAAC 5760  
AATTTATTAT TGATCTCTGA ATCTGTAAGA TCAATAACA AGTCTCTATC CATGTTACCA 5820  
AATTTAACTT TTTGAAATA ATAACTTTA AATATCAGA TGTGTTATTA CAGGATGATA 5880  
CTTGAATCA AGTGAATGA GTTATATGGT CATCACTAAA TTAGAAATC TATTGTGAAA 5940  
CAAGACAAA CAGGAAAGTA CAGAATAGAG ACTTTTAGTA AATAATGGA ATTTAAAGA 6000  
AAGTGTATT TTACAGTGTG ACGACAGAAA AGGATGTCTT TGTGTGATA GTCTTTGAGG 6060  
GATCTCCGTA AATCTGCGG CACAGGTACA AGAAATAGCC AATATTAGT TCCAGACCA 6120  
TGTTTAGTAG TGTCCAGTT CAGATCATGC TGCCAAGAGG TATCTCCCCC TCAGTGGGT 6180  
CATCACTGAG CCTGGAATT GGAGACTCAT ACTTGCCAG CACAATGTTA CGGCAGACA 6240  
GGCCGACATC TATGATTAGC TAGAAGCCAT AAGAAAAGC TGCTAAGTGG CCACATAGGT 6300  
CCACTTTCT GTTTTGTAA TGCTTTTATT AGCAGATCTT TTTTTCCTA GCTCCATGG 6360  
GCCTATGAGA GGCATTATG ATTTTGTGTC CTACAATAAG TCAGCTGTC TGGTGTGAGT 6420  
TGTTTATGA GAAATGCTTT CCAAGGAGG TCTAGGAAGA TCTGACACA TAAGAACTTT 6480  
GGCTTAGAGA GCTTCCAGG TGTAGTGCCA ATAAAACTG ACCTGGAAG AAAACCTGCC 6540  
CAGCAGGAA CATGCTTCT GAACTCACT GAGAGTGTAT GGTGTATGTC ACTTCTCATA 6600  
TATTCTTAG TTTGATTG TCTTTTATC AATTTTAGC TCTTTTCCAG TTCACTTGTG 6660  
CTGCTCTGA TATTGGTATT TTTAAATTT TGTGTAAT AATGAAAGA GTGAAATAT 6720  
ATTTTATAAT TACTCATTTG TAGTTTTTTT TTTAATTTA ATAACTTCC TCCAAAAGT 6780  
GCTCCCTTAA AA

Seq ID NO: 166 Protein sequence:  
Protein Accession #: AAG34652

1 11 21 31 41 51  
MAAEBAAGG KVLREENQC IAPVSSSRVS PGRPTAMGS FSSHMTPEPR KRKGSDDPS 60  
QVEDGEHVQV MKAFFRAHSQ TKRRRDKMN NLIELSAMI PQCNPMARKL DKLTVLRMAV 120  
QHLRLKGLT NSYVGSNYRP SPLQDNELRH LILKTAEGFL FVVGCEGKI LFKVSKSVKI 180  
LNYDAQSLTG QSLDFDLHPK DVAKVKEQLS SFDISPREKL IDAKYGLQVH SNLHAGRTRV 240  
YSGSRSPFC RKSKCKISVK ESHGCLPNSK KKEHRKFYTI HCTGYLRSWP PNIVGMEER 300  
NSKKNNSFT CLVATGLQRP YIVPQNSGEI NVKPTFPIR FAVNGKPVYV DQRATALLGY 360  
LPQELLGTS YEPFQDDHN NLTDKHKAVL QSKKILIDS YKFRAGDGSF VTLKQWFSF 420  
TNPWTKLEY IVSVNTLVLG HSEPGASFL PCSSQSSEES SRQSCMSVPG MSTGTVLGAG 480  
SIGTDIANEI LDQLRLQSS YLDDSSPTGL MKDTHVNCR SMSNKLFPF SPSEMGLEA 540  
TRQNQSTVAV HSHEPLLSDG AQLDFDALCD NDDTAMAAFM NYLEAEGGLG DPGDFSDIQW 600  
TL

Seq ID NO: 167 DNA sequence  
Nucleic Acid Accession #: NM\_014400  
Coding sequence: 86-1126

1 11 21 31 41 51  
GGTTACTCAT CTTGGGCTCA GTTAAGAGGG CCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60  
GAGCCCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAGG AAAGCAGGTG CCCAGGCCAT 120  
GATCTGGAAT GCAGGCTGGC TGCTGCTGCT GCTGCTTGGC GGAGGAGCGC AGGCCCTGGA 180  
GTGCTACAGC TGCGTGACAGA AAGCAGATGA CGGATGCTCC CCGAACAGA TGAAGACAGT 240  
GAAGTGCGCG CCGGCGGTGG ACCTCTGCAC CGAGGCGGTG GGGGCGGTGG AGACCATCCA 300  
CGGACAATTC TCGCTGGCAG TGCSGGGTG CGGTTGGGA CTCGCCGCA AGAATGACCG 360  
CGGCTCGGAT CTTCAAGGCG TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420  
CTGCAAGGCC AAGCTCAACC TCACCTCGCG GCGCTCGAC CCGGCAGGTA ATGAGAGTGC 480  
ATACCCGCCC AACCGGCTGG AGTGCTACAG CTGTGTGGCG CTGAGCCGGG AGGCGTGCCA 540  
GGGTACATCG CCGCGGTGCG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGCGTG 600  
CTTCGACGGC AAGCTCACTT TGACGGCAGC TAATGTGACT GTGTCTTGTG CTGTCCGGGG 660  
CTGTGTCCAG GATGAATCT GCACCTCGGA TGGAGTAACA GGCCAGGGT TCAAGCTCAG 720  
TGCTCTCTGT TGCCAGGGGT CCGCTGTAA CTCTGACCTC CGCAACAAGA CTAATTCTC 780  
CCCTCGAATC CCACCCCTTG TCGGCTGCC CCCTCCAGAG CCCAGACTG TGGCTCAAC 840  
CACATCTGTC ACCACTCTTA CTTGGGCCCC AGTGAGACCC ACATCCACCA CCAACCCAT 900  
GCCAGCGCCA ACCAGTCAGA CTCCGAGACA GGGAGTAGAA CACGAGGCTT CCCGGATGA 960  
GGAGCCGAGG TTGACTGAGG GCGCGCTGG CCACAGGAC CGCAGCAAT CAGGGCAGTA 1020  
TCCTGCAAAA GGGGGGCCCC AGCAGCCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080  
ATTGGCAGCC CTCTGTGTGG CCGTGGCTGC TGGTGTCTTA CTGTGAGCTT CTCACCTGG 1140  
AAATTTCCCT CTCACCTACT TCTCTGGCCC TGGGTACCCC TCTTCTCATC ACTTCTGTT 1200  
CCCAACCTG GACTGGGCTG GCCCAGCCCC TGTTTTCCA ACATTCGCCA GTATCCGAG 1260  
CTCTGCTGTC GCTGTTTTCG GCCTTTGGGA AATAAAATAC CGTTGTATAT ATTCTGGCAG 1320  
GGGTGTTCTA GCTTTTGTAG GACAGCTCCT GTATCCTTCT CATCCTGTG TCTCCGCTTG 1380  
TCTCTTGTG ATGTTAGGAC AGAGTGAGAG AAGTCAGCTG TCAAGGGGAA GGTGAGAGAG 1440  
AGGATGCTAA GCTTCTACT CACTTTCTCC TAGCCAGCCT GGACTTTGGA GGTGTGGGTG 1500  
GTTGGGACAA GCTTCCCCA CTCTAAGCAC TGCCCTCCCT ACTCCCGCA TCTTTGGGGA 1560  
ATCGGTTCCT CATATGCTT CTTACTAGA CTGTGAGCTC CTGAGGGCA GGGACCGTGC 1620  
CTTATGTCTG TGTGTGATCA GTTCTGCGCA CATAAATGCC TCAATAAAGA TTTAATTACT 1680

WO 02/086443  
TTGTATAGTG AAAAAAAA

PCT/US02/12476

Seq ID NO: 168 Protein sequence:  
Protein Accession #: NP\_055215

```
1      11      21      31      41      51
|      |      |      |      |      |
MDPARKAGAQ AMIWTAGWLL LLLLRRGGAQA LECYSCVQKA DDGCSNPKMK TVKCAPGVDPV 60
CTEAVGAVET IHGQFSLAVX GCGSGLPGKN DRGLDLHGLL AFIQLQCCAQ DRCNAKLNLT 120
SRALDPAGNE SAYFPNGVEC YSCVGLSREA CQGTSPVVS CYNASDHVYK GCFDGNVTLT 180
AANVTVSLPV RGCVDDEFCT RDGVTGPGPT LSGSCCQGSR CNSDLRNKTY FSPRIPLPLVR 240
LPPPEPTTVA STTSVTSTST APVVRTSTTK PMPAPTSQTP RQGVHEASR DEEPRLTGGA 300
AGHQDRSNSG QYPAKGSPQO PHNKGCVAPT AGLAALLLAV AAGVLL
```

Seq ID NO: 169 DNA sequence  
Nucleic Acid Accession #: NM\_006875  
Coding sequence: 186-1190

```
1      11      21      31      41      51
|      |      |      |      |      |
GAATTCGGCA CGAGCGCGCG GCGAATCTCA ACGCTGCGCC GTCTGCGGGC GCTTCGGGGC 60
CACCAGTTTC TCTGCTTTCC ACCCTGGGCG CCCCAGCCCC TGGCTCCCCA GCTGCGCTGC 120
CCCGGGCGCT CACGCCCTGC GGGCTTAGCG GGTTCAGTGG GCTCAATCTG GCGAGCGCCA 180
CTCCCATGTT GACCAAGCCT CTACAGGGCG CTCGCGCGCC CCGCGGGACC CCAACGCGCG 240
CGCCAGGAGG CAAGGATCGG GAAGCGTTGG AGGCGAGTA TCGACTCGCG CCCCTCCTGG 300
GTAAGGGGGG CTTTGGCACC GTCTTGCAG GACACCGCCT CACAGATCGA CTCGAGGTGG 360
CCATCAAAAGT GATTCCCCGG AATCGTGTGC TGGGCTGGTC CCCCTTGTCA GACTCACTCA 420
CATGCCCACT CGAAGTCGCA CTGCTATGGA AAGTGGGTGC AGGTGGTGGG CACCTTGGCG 480
TGATCCGCTT GCTTGACTGG TTTGAGACAC AGGAAGGCTT CATGCTGGTC CTCGAGCGCG 540
CTTTGCCCGC CCAGGATCTC TTTGACTATA TCACAGAGAA GGGCCCACTG GGTGAAGGCC 600
CAAGCGCTGT CTTCTTGGC CAAGTAGTGG CAGCCATCCA GCACTGCCAT TCCCGTGGAG 660
TTGTCCATCG TGACATCAAG GATGAGAACA TCCTGATAGA CCTACGCGGT GGCTGTGCCA 720
AACTCATGGA TTTTGGTTC GTGCGCTGCT TTCATGATGA ACCCTACACT GACTTTGATG 780
GGACAAGGGT GTACAGCCCC CCAGAGTGGG TCTCTCGACA CCAGTACCAT GCATCCCGG 840
CCACTGTCTG GTCACTGGGC ATCCTCCTCT ATGACATGTT GTGTGGGGAC ATTCCCTTTG 900
AGAGGGACCA GGAGATTCTG GAAGCTGAGC TCCAATTCCC AGCCCATGTC TCCCGAGACT 960
GCTGTGCCCT AATCCGCGCG TGCTGGCGCC CCAAACTTCT TTCCGAGCCC TCACTGGAAG 1020
AGATCCTGCT GGACCCCTGG ATGCAAAAC CAGCGAGGGA TGTACCCCTT CAACCCCTCC 1080
AAAGGAGGCC CTGCCCCCTT GGCCTGGTCC TTGCTACCCT AAGCCTGGCC TGGCTGGGCC 1140
TGGCCCCCAA TGGTCAGAAG AGCCATCCCA TGGCCATGTC ACAGGGATAG ATGGACATT 1200
TTGACTTGG TTTTACAGGT CATTACAGT CATTAAAGTC CAGTATTACT AAGGTAAGGG 1260
ATTGAGGATC AGGGGTTAGA AGACATAAAC CAAGTTTGCC CAGTTCCTTT CCAATCTCTA 1320
CAAAGGAGCC TTCTCCCAAT AACCTGTGGT CCCTGATTTT GGAGGGGGAA CTCTGTGCTT 1380
CTCATTTTGC TAAGGAAGTT TATTTTGGTG AAGTTGTTCC CATTTTGAGC CCCGGGACTC 1440
TTATTTTGTG GATGTGTAC CCCACATTG CACTCCTTAC TACCACCACA CAAACTTAGT 1500
TCATATGCTT TTACTTGGGC AAGGGTGTCT TCCTTCCAAT ACCCCAGTAG CTTTATTATT 1560
AGTAAAGGGA CCCTTTCCCC TAGCCTAGGG TCCCATATTG GGTCAAGCTG CTTACCTGCC 1620
TCAGCCGAGG ATTTTATTAT TTGGGGGAGG TAATGCGCTG TTGTTACCCC AAGGCTTCTT 1680
TTTTTTTTTT TTTTTTTTGG GGTGAGGGGA CCTACTTTTG TTATCCCAAG TGCTCTTATT 1740
CTGGTGAGAA GAACCTTAAT TCCATAATTT GGAAGGAAT GGAAGATGGA CACCACCGGA 1800
CACCACCAAG CAATAGGATG GGATGGATGG TTTTGTGGGG GATGGGCTAG GGAATAAAG 1860
GCTTGTCTGT TGTTTTCCCT GGGCGCTCCC TCCAATTTTG CAGATTTTTG CAACCTCCTC 1920
CTGAGCCGGG ATTGTCCCAT TACTAAATG TAAATAATCA CGTATTGTGG GGAGGGGAGT 1980
TCCAAGTGTG CCTCCTTTT TTTTCTGCTG TGAATTATTT AAAAAGCCAT GTGTGGAAC 2040
CCACTATTTA ATAAAGTAA TAGAATCAGA AAAAAA AAAAAA
```

Seq ID NO: 170 Protein sequence:  
Protein Accession #: NP\_006866

```
1      11      21      31      41      51
|      |      |      |      |      |
MLTKPLQGPP APPGTPTPPP GGDREAFEA EYRLGPLLKK GPGPTVFAGH RLTDRLQVAI 60
KVIPRNRVLG WSPILSDSVTC PLEVALLWKV GAGGGEHGVY RLLDWFTQE GPMLVLERPL 120
PAQDLFDYIT EKGPLGEGPS RCFEGQVVA IQHCHSRGVV HRDIKDENIL IDLRRGCAKL 180
IDFGSGALLH DEPYTDFDGT RVYSPPEWIS RHQYHALPAT VWSLGLILYD MUCGDIFFER 240
DQBILEAELH PPARHSPDCC ALIRRCIAPK PSSRPSLEBI LLDPMNQTPA EDVTPQPIQR 300
RPCFFGLVLA TSLAWPGLA PNGKSHPMA MSQG
```

Seq ID NO: 171 DNA sequence  
Nucleic Acid Accession #: NM\_003646  
Coding sequence: 89..2875

```
1      11      21      31      41      51
|      |      |      |      |      |
GCGGCGCGGA GCGGGCGTGT TGAGCCCGGG CGCGCGGCCC GGCATGGGCG TCTCCCGGG 60
GCCTCCGCGC GCGCGGGGCT AGGGCGGAT GGAGCCGCGG GACGGTAGCC CCGAGCCCG 120
GAGCAGCGAC TCCGAGTCGG CTTCCGCTCT GTCCAGCGCG TCCGAGCGCG ACGCGGCTCC 180
CGAGCCGGAC AAGGCGCGCG GCGGACTCAA CAAGCGGCGC TTCCGCGGGC TCGCGCTCTT 240
CGGGCACAGG AAAGCCATCA CCAAGTCGGG CCTCCAGCAC CTGGCCCCCC CTCGCCCCAC 300
CCCTGGGGCC CCGTGCAGCG AGTCAGAGCG GCAGATCCGG AGTACAGTGG ACTGGAGCGA 360
GTCAAGGACA TATGGGGAGC ACATCTGGTT CGAGACCAAC GTGTCCGGGG ACTTCTGCTA 420
CGTTGGGGAG CAGTACTGTG TAGCCAGGAT GCTGAAGTCA GTGTCTGAA GAAAGTGCSC 480
AGCCTGCAAG ATTGTGGTGC ACACGCCCTG CATCGAGCAG CTGGAGAAGA TAAATTTCCG 540
CTGTAAGCCG TCCTTCCGTG AATCAGGCTC CAGGAATGTC CGGAGGCCAA CCTTTGTACG 600
GCACCACTGG GTACACAGAC GAGCCAGGA CGGCAAGTGT CGGCACTGTG GGAAGGGATT 660
CCAGCAGAAG TTCACCTTCC ACAGCAAGGA GATTGTGGCC ATCAGCTGCT CGTGGTGCAA 720
```

GCAGGCATAC CACAGCAAGG TGCTCTGCTT CATGCTGCAG CAGATCGAGG AGCGGTGCTC 780  
 GCTGGGGGTC CACGCGCCG TGCTCATCCC GCCACCTGG ATCCTCCGGG CCGCGAGGCC 840  
 CCAGAACTACT CTGAAGACAA GCAAGAAGAA GAAGAGGGCA TCCTTCAAGA GGAAGTCCAG 900  
 CAAAGAAAGG CTTGAGGAGG GCGCTGGAG ACCCTTCATC ATCAGGCCCA CCCCCTCCCC 960  
 5 GCTCATGAAG CCCCTGCTGG TGTTTGTGAA CCCCAGAGT GGGGGCAACC AGGGTGCAAA 1020  
 GATCATCCAG TCTTTCCTCT GGTATCTCAA TCCCAGCAA GTCTTCGACC TGAGCCAGGG 1080  
 AGGGCCCAAG GAGGCGCTGG AGATGTACCG CAAAGTGAC AACCTGCGGA TCCTGGCGTG 1140  
 CGGGGGCGAC GGCACGCTGG GCTGGATCCT CTCACCCCTG GACCAGCTAC GCCTGAAGCC 1200  
 10 GCCACCCCTT GTTGCCATCC TGCCCTGGG TACTGGCAAC GACTTGGCCC GAACCTTCAA 1260  
 CTGGGGTGGG GGTACACAG ATGAGCCTGT GTCCAGATC CTCTCCACG TGGAGGAGGG 1320  
 GAACGTGGTA CAGCTGGACC GCTGGACCT CCACGCTGAG CCAACCCCG AGGCAGGGCC 1380  
 TGAGGACCGA GATGAAGGCG CCACCGACCG GTTGCCCTG GATGTCTCA ACACTACTT 1440  
 CAGCCTGGGC TTGACGCCC ACGTACCCCT GGAGTTCAC GAGTCTCAG AGGCCAACCC 1500  
 15 AGAGAAATTC AACACGCGCT TTGGAAATAA GATGTCTAC GCGGGACAG CTTTCTCTGA 1560  
 CTTCTGATG GGCAGCTCCA AGGACCTGGC CAAGCACATC CGAGTGGTGT GTGATGGAAT 1620  
 GGACTTGACT CCAAGATATC AGGACCTGAA ACCCAAGTGT GTTGTTTTCC TGAAACATCCC 1680  
 CAGGTACTGT GCGGCCACCA TGCCCTGGGG CCACCTGGG GAGCACCACG ACTTTGAGCC 1740  
 CCAGCGCAT GACGACGGCT ACCTCGAGGT CATTGGCTTC ACCATGAGT CGTTGGCCGC 1800  
 20 GCTGCAGGTG GCGGACACG GCGAGCGCT GACGAGTGT CCGAGGTGG TGCTCACCAC 1860  
 ATCCAGGCC ATCCCGGTGC AGGTGGATGG CGAGCCCTGC AAGCTTGAC CCTCAACGAT 1920  
 CGCATCGCC CTGCGCAACC AGGCCACCAT GGTGCAGAA GCGAAGCGCG GGAGCGCGCG 1980  
 CCCCCTGCAC AGCGACACG AGCGGTGCC AGAGCAGTTG CGCATCCAGG TGAGTCGCGT 2040  
 CAGCATGCAC GACTATGAGG CCTGCACTA CGACAAGGAG CAGCTCAAGG AGGCCTCTGT 2100  
 25 GCGCTGGGC ACTGTGGTGG TCCAGGAGA CAGTGACCTA GAGCTCTGCC GTGCCACAT 2160  
 TGAGAGACTC CAGCAGGAGC CCGATGTGTC TGGAGCCAG TCCCAGCAT GCCAGAACT 2220  
 GTCCCCAGG TGGTGTCTTC TGGACGCCAC CACTGCCAGC CGCTTCTACA GGATCGACCG 2280  
 AGCCAGGAG CACCTCAACT ATGTGACTGA GATGCACAG GATGAGATT ATATCCTGGA 2340  
 CCCTGAGCTG CTGGGGCATC CGCCCGGCC TGAACCTCCA ACCCCACTT CCCCCTCTCC 2400  
 30 CACCTCACCC TGCTCACCCA CGCCCGGTC ACTGCAAGG GATGCTGCAC CCCCTCAAGG 2460  
 TGAAGAGCTG ATTGAGGCTG CCAAGAGGAA CGACTTCTGT AAGCTCCAGG AGCTGCACCG 2520  
 AGCTGGGGGC GACCTCATGC ACCGAGACGA CGAGAGTGC ACCTCCCTG ACCACGAGT 2580  
 CAGCACTGGC AGCAAGGATG TGGTCCGCTA CTGCTGACG CACGCCCCCG CAGAGATCCT 2640  
 35 TGATGCGGTG GAGGAAACG GGGAGACCTG TTTGCACCA GCAGCGGCC TGGGCCAGCG 2700  
 CACCATCTGC CACTACATCG TGGAGGCGCG GGCCTCGCTC ATGAAGACAG ACCAGCAGGG 2760  
 CGCACTCCCC CGCAGCGGG CTGAGAAGGC TCAGGACACC GAGCTGGCGC CCTACCTGGA 2820  
 GAACCGGCAG CACTACAGA TGATCCAGCG GGAGGACACG GAGACGCGTG TGTAGCGGGC 2880

Seq ID NO: 172 Protein sequence:  
 Protein Accession #: NP\_003637

1 11 21 31 41 51  
 MEPRDGSPEA RSSDSESASA SSSGSEADAG PEPDKAPRRL NKRRFPGLRL FGRKAITKS 60  
 45 GLQLHAPPPP TPQAPCSESE RQIRSTVDMS ESATYGEHIW FETNVSGDFC YVGEQYCVAR 120  
 MLKSVSRKRC AACKIVVHTP CIEQLEKINF RCKPSFRESG SRNVREPTFV RHHVHRRRQ 180  
 DGKCRHRCGK PQQKTPFRSK EIVAISCSNC KQAYHSKVSC PMLQIIEPC SLGVHAHVVI 240  
 PPTWILRARR PONTLKASKK KKRASFRRKS SKKGPEEGRW RPFIIIRTPS PLMKPLLVFV 300  
 NPKSGGNQGA KIIQSFNLVL NPROVFDLSQ GGPKEALEMY RKVHNLRLLA CGGDGTVGWI 360  
 50 LSTLDQLRLK PPPFVAIIFL GTGNDLARTL NNGGYYTDEP VSKILSHVEE GNVVQLDRWD 420  
 LHAEPNEPAG PEDRDEGATD RLPLDVFNNY FSLGPDABVT LEFHESREAN PEKFNRSFRN 480  
 KMFYAGTAPS DPLMGSSKDL AKHIRVVC DG MDLTPKIQDL KPQCVFLNI PRYCACTMPW 540  
 GHFGEHHDPE PQRHDDGYLE VIGPTMTSLA ALQVGGHGER LTQCREVVL TSKAIPVQVD 600  
 GEPCKLAASR IRIALRNQAT MVQKAKRRA APLHSDQQPV PEQLRIQVSR VSMHDYEAHL 660  
 55 YDKBQLKEAS VPLGTVVVPG DSDLELCRAH IERLQOEPDG AGAKSPTQCK LSPKWCFLDA 720  
 TTSARFYRID RAQEHNLVYT EIAQDEIYIL DPELLGASAR PDLPFTSPL PTSPCSPTPR 780  
 SLQGDAAFPQ GBELIEAAKR NDFCKLQELH RAGGDLMRD EQSRTLLHHA VSTGSKDVVR 840  
 YLLDHAPPEI LDAVENGET CLHQAAALGQ RTICYIYVEA GASLMKTDQQ GDTPRQRAEK 900  
 AQDTLAAVL ENRQHYMIQ REDQETAV

Seq ID NO: 173 DNA sequence  
 Nucleic Acid Accession #: AF232772  
 Coding sequence: 1-1662

1 11 21 31 41 51  
 65 ATGCCCGTGC AGCTGACGAC AGCCCTGCGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60  
 GTGCTGGGTG GCATCTGGC AGCCTATGT ACGGGCTACC AGTTCATCCA CACGGAAAAG 120  
 CACTACCTGT CCTTCGGCT GTACGGCGCC ATCCTGGGCC TGCACTGTCT CATTGAGAGC 180  
 70 CTTTTCGCT TCCTGGAGCA CCGGCGCATG CGACGTGCGG GCCAGGCCCT GAAGCTGCCC 240  
 TCCCCTGGG GGGGCTCGGT GGCACGTGTC ATTGCGCAT ACCAGGAGGA CCCTGACTAC 300  
 TTGCGCAAGT GCCTGCGCTC GCGCCAGGCG ATCTCCTTCC CTGACCTCAA GGTGGTCATG 360  
 GTGGTGGATG GCAACCGCCA GGAGGACGCC TACATGTCTG ACATCTTCCA CGAGGTGCTG 420  
 75 GCGCGCACCG AGCAGGCGCG CTTCTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGCGAG 480  
 GGTGAGACGG AGGCAGCCT GCAGGAGGGC ATGGACCCTG TCGGGGATGT GGTGCGGGCC 540  
 AGCACCTTCT CGTGATCATL GCAGAAGTGG GGAGGCAAGC GCGAGGTGAT GTACACGGCC 600  
 TTCAAGGCCC TCGGCGATTG GTTGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660  
 GATCCAGCCT GCACCATCGA GATGCTTCCA GTCTGGAGG AGGATCCCCA AGTAGGGGGA 720  
 GTGCGGGGAG ATGTCCAGAT CCTACACAAG TACGACTCAT GGATTTCCTT CCTGAGCAGC 780  
 80 GTGCGGTACT GGATGGCCTT CAAAGTGGAG CGGGCCTGCC AGTCTACTT TGGCTGTGTG 840  
 CAGTGTATTA TGGTCCGCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCTGGAG 900  
 GACTGTGACC ATCAGAAGTT CCTAGGCAGC AAGTGCACT TCGGGGATGA CCGGCACCTC 960  
 85 ACCAACCGAG TCCTGAGCCT TGGCTACCGA ACTAAGTATA CCGCGCGCTC CAGTGCCTC 1020  
 ACAGAGACCC CACTAAGTA CTTCCGGTGG CTCACACGAG AACCCCGCTG GAGCAAGTCT 1080  
 TACTTCCGGG AGTGCTCTA CAACTCTCTG TGGTTCCTA AGCACCACCT CTGGATGACC 1140  
 TACGAGTCAG TGGTCAAGGG TTCTTCCCTC TTGCAACGCT TATACAGCTT 1200  
 TTCTACCGGG GCGCATCTG GAACATCTC CTCTTCTGCG TGAAGGTGCA GCTGGTGGG 1260  
 ATTATCAAGG CCACCTACGC CTGCTTCTT CCGGCAATG CAGAGATGAT CTTTATGTCC 1320

CTCTACTCCC TCCTCTATAT GTCCAGCCTT CTGCCGCCA AGATCTTTGC CATTGCTACC 1380  
 ATCAACAAAT CTGGCTGGGG CACCTCTGGC CGAAAAACCA TTGTGGTGAA CTTCACTGGC 1440  
 CTCATTCCTG TGTCCATCTG GGTGGCAGTT CTCTGGAGG GGCTGGCCTA CACAGCTTAT 1500  
 TGGCAGGACC TGTTCACTGA GACAGAGCTA GCCTTCCTTG TCTCTGGGGC TATACTGTAT 1560  
 GGCTGCTACT GGGTGGCCCT CCTCATGCTA TATCTGGCCA TCATCGCCCG GCGATGTGGG 1620  
 AAGAAGCCCG AGCAGTACAG CTTGGCTTTT GCTGAGGTGT GACATGGCCC CCAAGCAGAG 1680  
 CGGTTAAAGT GCAATGGGTA AGGGAGGGAA GGGGAATGGA AGAGAAAAGA CAGGGTGGGA 1740  
 GGGAGGAGGG AGTGCTGTGT TTTAGTCTCT TAATGGTCCA AAGGACAAAT CTAAAATGCA 1800  
 AAGAACGGTG ATGTAGTAGT GCCTGACAGC TCTGTTTAGA GGAGGCAACA CTGATCCCCC 1860  
 AGATGCAGGG CTGCAGGGGA TTCTGTGTTT TCAGACTGCC TGTCTGCTTG CATCTGCACA 1920  
 TAGGCAGTAG CCTCTCTCTG GGCTCCAGAG GGCACCTAGA AGTTGTGCTA AACCAAGTTA 1980  
 AGTCCCATTC AGTGGCAACT TGTGATAGGT ACCTGAGTGA CGGCAACCTG CGGAAGGAGG 2040  
 TTCTCCAGCG CCATCTGAAC ACAACCAAGG GTGGCAGGAG AATTTCTACT GAGCGAGGTG 2100  
 GGCGGCTTAG TGATGTGCTA CCCCACCCCA CCCATAAGTA GTCATCAATG CAATAAGATT 2160  
 GGGGTGAGA TACAGAGCCC AGAAGCCTGA TCTTTGGGCA TCAGAAAACA GGGTCCAGGA 2220  
 ATGTGTCTTT ATGTGAGATA CCCCACCTCA CATCAACATT CCAGGGATGA GCCAAACAG 2280  
 CAGGGAGTTA GCACTGAAC TCTTTTAAAA GTGCACATTA AAAAGGAAG TTTGCCAGGA 2340  
 GGAACAAGA GATTGTGGTG GTGCTAAAGG AGGCCATAAG CTACACAGAG GCCTTGGGTG 2400  
 TTCCACCTGG AAACCTGGTA GACGTCTAGA TGGGTTCTTA GCTTGTCTGT GATCTCTGCT 2460  
 GGGGAGATAA AAGATTAAAG CCCCAACATG TTCAGAAAAG AAGTGAAGTC TTGGGTATTT 2520  
 TAACCTGTAT ACTCTGAAT TCTCTCAAA TTCAGCTCTG ATCTGAGGCT AAGACACACT 2580  
 CCCCACCTCA CTTTCTTCAA AGCCACATTT TTTGAGGTAT CACTGCAGTC ACCTCTTCTA 2640  
 CCTCATCAT CATAGGTAAG GTTTTCAAGG TGGCAATTGG GCGGAGGCC CGGCTTCTTA 2700  
 TAGAAGCTTC AGCAGGAGGC AAGCGTGTTC TCAGCACATA TGGGAACAT GAGGAGCCTC 2760  
 TGATCAAAAT GGCTACAATC TTGGAGCTGC TTGGAGCGAT TCCTTGGCAG CCGGGTTAGC 2820  
 ATGTGTGACT TTCAGGCTAC TGTTCTTGAC AATCATCTCC AATGGAAGC TTTTCACTGT 2880  
 TCCCAAGGTG AACTCTCAA TCCAAATGG TTATCTTGA GACCATCCAT TCTCCTCAGT 2940  
 GGCTTCTCCA GGAATTCTT ACAGCCAAGT TGTGACAGTC ACTGCATTTG CCTGCTTCTT 3000  
 TCCAGAAACC AACTAGGAG ATGAACTGG TTCTACATC CTAAAGTTCT TGCTTTCTCT 3060  
 CTCATGCTCT CTGAGGCTGT TTTTGGCTGT TTTCCCTCTG CTGCTTTTGG GGAATGAGGG 3120  
 GAAGCCATT TCCAAGTACT TTGCAATCCA GCGTGTCTC AGCGTTTGA GTTTAAACCC 3180  
 TGGGATCCTG ACTAAGCCTT TGACTTAAGG GTTGCTTGCT TGCCCTCCAA ATGTCTTTC 3240  
 TCAAAGGGGC CACTAACCCT GTGCAGAAC AGCCTAAGG TGGACAGCAG ACAAGAGGGC 3300  
 AAGCCTCTAA TGTACCAAGT GCTTCTTACA AAGAAGCAAG GTGTGCTCCG AACCAAGAT 3360  
 GGGCAAAACC TGGTGTCTT CTTTATCTCC CACGAACCTA AGGGTTTCC AAGTGTAGCT 3420  
 AACAGTTGCC ACATCACACA GACCTCCAGT TTCTGGTAAG ACTGCTGGTT GACATCAGAC 3480  
 CCAACCCATT GAAGGCTGGA AGGCAGCAGG CATTGTCTAA GGCAGCTGAT CCAGGCAATC 3540  
 GTTCTGCTGG CCAAGAAGTT AAATATTTT GAGCATTAGA ATGAGGAAA TCCGCTCAGC 3600  
 CAAGTGACAGA TGTGACACT GCTTAAGGGC TTGTTTTCT TCAGCATTTA CTTGAGATT 3660  
 AATGTAGGAT GACAGGCTCT CCTGGCTGTC CTACCATCAG CTCTGCTTG CACTGTGGTC 3720  
 GTCAACTTTC CTCAAATCAA AAACAGGCAG GTACAGTAG TGGGCTCACA ACCTTTGACC 3780  
 TCGACTGGTT TTTCTAAGT ATTTTGTACA TTTTTCAGCA GCAAAACCAA ACTGGGCTCT 3840  
 CAGCTTATC CCGCTTTCTT GCAAGGGAAG AGCCTTTATA CAATTGGAGC CATTTTGGTT 3900  
 TTTCTCAT TGAATTTCAA ATCTCTTTT GTATTGTTT TACAATAAT TGTAAACATA 3960  
 TTTATTTTA CTGCTTTT TTTTTTTTT TAATTTTTCAG GTCAAGTTT TTATACTGCA 4020  
 CTTATTTTGA AAAATAAGA TTCTACAT

Seq ID NO: 174 Protein sequence:  
 Protein Accession #: AAF36984

1 11 21 31 41 51  
 MPVQLTALR VVGTSLFALA VLGGILAAVY TGYQFIHTEK HYLSPGLYGA ILGLHLIIQS 60  
 LFAFLERHRM RRAGQALKLP SPRRGSVLAL IAAQEDPDY LRKCLRSAGR ISFPDLKVV 120  
 VVDGNRQEDA YMLDIFHEVL GGTEQAGFFV WRNPFHEAGE GETEASLQEG MDRVRDVVRA 180  
 STFSIMQKQW GGRREVMTYA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEEDPVQVG 240  
 VGGDVQLINK YDSWISFLSS VRYWMAFNVE RACQSYFGCV QCISGPLQMY RNSLLQQFLE 300  
 DWYHQKFLGS KCSFGDDRHL TNRVLSLGYR TKYTARESKL TETPTKYLRLV LNQOTRWSKS 360  
 YPREWLINSL WFKHHLWMT YESVVTGFPF PFLIATVQL FYRGRIWNIL LFLLTQLVLG 420  
 IIAKYACFL RGNAEMIFMS LYSLLYMSL LPAKIPAIAT INKSGNWTSG RKTIVNVPFG 480  
 LIPVSIWVAV LLEGLAYTAY QDLFSETEL AFLVSGAILY GCYVWALLML YLAIIRRCG 540  
 KPBPQYSLAF AEV

Seq ID NO: 175 DNA sequence  
 Nucleic Acid Accession #: NM\_000691  
 Coding sequence: 43..1404

1 11 21 31 41 51  
 CCAGGAGCCC CAGTTACCGG GAGAGGCTGT GTCAAAGCGG CCATGAGCAA GATCAGCGAG 60  
 GCGGTGAAGC GCGCCCGCGC CGCCTTCAGC TCGGGCAGGA CCCGTCCGCT GCAGTTCGGA 120  
 TTCCAGCAGC TGGAGGCGCT GCAGCGCCTG ATCCAGGAGC AGGAGCAGGA GCTGGTGGGC 180  
 GCGCTGCGCG CAGACCTGCA CAAGAATGAA TGGAAACGCT ACTATGAGGA GGTGGTGTAC 240  
 GTCCTAGAGG AGATCGAGTA CATGATCCAG AAGCTCCCTG AGTGGGCGCG GGATGAGCCC 300  
 GTGGAGAGA GCGCCAGAC TCAGCAGGAC GAGCTCTACA TCCACTCGGA GCCACTGGGC 360  
 GTGGTCTCTG TCATTGGCAC CTGGAAGTAC CCCTTCAACC TCACCATCCA GCCATGGTG 420  
 GCGGCCATCG CTGCAGGGAA CGCAGTGTTC CTCAAGCCCT CGGAGCTGAG TGAGAACATG 480  
 GCGAGCCTGC TGGCTACCAT CATCCCCAG TACCTGGACA AGGATCTGTA CCCAGTAATC 540  
 AATGGGGGTG TCCCTGAGAC CACGGAGCTG CTCAAGGAGA GGTTCGACCA TATCCTGTAC 600  
 ACGGGCAGCA CGGGGGTGGG GAAGATCATC ATGACGGCTG CTGCAAGCA CCTGACCCCT 660  
 GTCAAGCTGG AGCTGGGAGG GAAGAGTCCC TGCTACGTGG ACAAGAACTG TGACCTGGAC 720  
 GTGGCCTGCC GACGCATCGC CTGGGGGAAA TTCATGAACA GTGGCCAGAC CTGCGTGGCC 780  
 CCAGACTACA TCCCTCTGTA CCCCTCGATC CAGAACCATA TTGTGGAGAA GCTCAAGAG 840  
 TCATCTGAAG AGTCTACCGG GGAAGATGCT AAGAAATCCC GGGACTATGG AAGAATCAT 900  
 AGTCCCGGCG ACTTCCAGAG GGTGATGGGC CTGATTGAGG GCCAGAGGT GGCCTATGGG 960  
 GGCACCGGGG ATGCGGCCAC TCGCTACATA GCCCCACCA TCCTCAGGGA CGTGGACCCC 1020

CAGTCCCGG TGATGCAAGA GGAGATCTTC GGGCCTGTGC TGCCCATCGT GTGCGTGCGC 1080  
 AGCCTGGAGG AGCCCATCCA GTTCATCAAC CAGCGTGAGA AGCCCTGGC CCTCTACATG 1140  
 TTCTCCAGCA ACGACAAGGT GATTAAAGAG ATGATTGCAG AGACATCCAG TGGTGGGGTG 1200  
 GCGGCCAAGC ATGTGATCGT CCACATCACC TTGCACTCTC TGCCCTCGG GGGCGTGGGG 1260  
 AACAGCGGCA TGGGATCTTA CCATGGCAAG AAGAGCTTCG AGACTTTCTC TCACCGCCGC 1320  
 TCTTGCTTGG TGAGGCTCTT GATGAATGAT GAAGGCTGA AGGTGAGATA CCCCCGAGC 1380  
 CCGGCCAAGA TGACCCAGCA CTGAGGAGGG GTTGCTCCGC CTGGCCTGGC CATACTGTGT 1440  
 CCGGCGGAG TGCGGACCAAC CCTCACTGGC TCTCTGGCC CTGGAGAATC GCTCTCGCAG 1500  
 CCGGCGGAG TGCGGACCAAC CCTCACTGGC TCTCTGGCC CTGGAGAATC GCTCTCGCAG 1560  
 GGGCCCGGCG CTACCATTC CAAGTCTCCA CCCCTTCTTA GACCAATAAA GAGACAAATA 1620  
 CAATTTCTA ACTCGG

Seq ID NO: 176 Protein sequence:

Protein Accession #: NP\_000682

1 11 21 31 41 51  
 MSKISEAVKR ARAAFSSGRT RPLQFRPQQL EALQRLIQEQ EQELVGALAA DLHKNEWNAY 60  
 YEEVVVLEE IEYMIQKLEP WAADEFVKT PQTQDELYI HSEPLGVVLV IGTWNYFFNL 120  
 TIQPMVGAIA AGNAVVLKPS ELSNEMASLL ATIIIPQYLDK DLYPVINGGV PETTELLKER 180  
 FDRILVTGST GVGKIIMTAA AKHLTPVLE LGGKSPCYVD KNCDLDVACR RIWKGKPMNS 240  
 GQTCVAPDYI LCDPSIQNI VEKLKSLKE FYGEDAKSR DYGRISARH FORVMGLIEG 300  
 QKVAYGGTGD AATRYIAPTI LTDVDPQSPV MQEEIFGPVL PIVCVRSLEE AIQFINQREK 360  
 PLALYMFSSN DKVIKMIAB TSSGGVAAND VIVHTLHSL PFGGVNGSGM GSYHGKKSFE 420  
 TFSHRRSLCV RPLMNDGLK VRYPPSPARM TQH

Seq ID NO: 177 DNA sequence

Nucleic Acid Accession #: NM\_001067.1

Coding sequence: 108-4703

1 11 21 31 41 51  
 CTAACCGAGC CGCTCTGTG GAGAAGCGGC TTGGTGGGG GTGGTCTCGT GGGGTCTTGC 60  
 CTGTTTAGTC GCTTTCAGGG TTCTTGAGCC CCTTCACGAC CGTCACCATG GAAGTGTCCAC 120  
 CATTGCGACC TGTAAATGAA AATATGCAAG TCAACAAAT AAAGAAAAAT GAAGATGCTA 180  
 AGAAAAGACT GTCCTGTGAA AGAATCTATC AAAAGAAAAC ACAATTGGAA CATATTTTGC 240  
 TCGGCCAGCA CACCTACATT GGTCTGTGG AATTAGTGAC CCAGCAAAATG TGGGTTTACG 300  
 ATGAAGATGT TGGCATTAACT TATAGGGAAG TCACCTTTGT TCCTGGTTTG TACAAAATCT 360  
 TTGATGAGAT TCTAGTTAAT GCTGCGGACA ACAACCAAG GACCCAAAA ATGCTCTGTA 420  
 TTAGAGTCAAC AATTGATCCG GAAAACAATT TAATTAGTAT ATGGAATAAT GAAAAAGSTA 480  
 TTCTGTGTGT TGAACACAAA GTTGAAAAAG TGTATGCCC AGCTCTCATA TTGGACAGC 540  
 TCCTAACTTC TAGTAACATG GATGATGATG AAAAGAAAGT GACAGGTGGT CGAAATGGCT 600  
 ATGGAGCCAA ATTGTGTAACT ATATTGAGTA CCAAAATTAC TGTGGAACA GCCAGTAGAG 660  
 AATACAGAAA AATGTTCAAA CAGACATGGA TGGATAATAT GGGAGAGCT GGTGAGATGG 720  
 AACTCAAGCC CTTCATGGA GAAGATTATA CATGTATCAC CTTCAGCCT GATTGTGCTA 780  
 AGTTTAAAT TGAAGCGCTG GACAAAGATA TTGTGCACT AATGGTCAGA AGAGCATATG 840  
 ATATTGCTGG ATCCACAAA GATGTCAAAG TCTTCTTAA TGGAAATAAA CTGCCAGTAA 900  
 AAGGATTCG TAGTATGTG GACATGTATT TGAAGGACAA GTTGATGAA ACTGGTAACT 960  
 CTTGAAAGT AATACATGAA CAAGTAAACC ACAGGTGGGA AGTGTGTTTA ACTATGAGTG 1020  
 AAAAAGGCTT TCAGCAAAAT AGCTTTGTCA ACAGCATTGC TACATCCAAG GGTGGCAGAC 1080  
 ATGTTGATTA TGTAGCTGAT CAGATTGTGA CTAACTTGT TGATGTTGTG AAGAAGAGA 1140  
 ACAAGGGTGG TGTTCAGTA AAAGCACATC AGGTGAAAA TCACATGTGG ATTTTGTGTA 1200  
 ATGCTTAAAT TGAACACCA ACCTTTGACT CTCAGACAAA AGAAAACATG ACTTTACAAC 1260  
 CCAAGAGCTT TGGATCAACA TGCCAATTGA GTGAAAAAT TATCAAAGCT GCCATTGGCT 1320  
 GTGGTATTGT AGAAAGCATA CTAACTGGG TGAAGTTTAA GSCCAAGTC CAGTTAAACA 1380  
 AGAAGTGTTC AGCTGTAAAA CATAATAGAA TCAAGGGAAT TCCCAAACTC GATGATGCCA 1440  
 ATGATGCGAG GGGCGAAGC TCCACTGAGT GTACGCTTAT CCTGACTGAG GGAGATTGAG 1500  
 CCAAACTTAT GGCCTGTTCA GGCCTTGGTG TGGTGGGAG AGACAAATAT GGGGTTTCC 1560  
 CTCTTAGAGG AAAAATACTC AATGTTGAG AAGCTTCTCA TAAGCAGATC ATGAAAAATG 1620  
 CTGAGATTAA CAATATCATC AAGATTGTGG GTCTTCAGTA CAAGAAAAAC TATGAGATG 1680  
 AAGATTCAAT GAAGACGCTT CGTTATGGGA AGATAATGAT TATGACAGAT CAGGACCAAG 1740  
 ATGGTTCCTA CATCAAGGC TTGCTGATTA ATTTATCCA TCACAATGG CCTCTCTTC 1800  
 TGGCAGATCG TTTCTGGAG GAATTATCA CTCCCATGT AAAGGTATCT AAAAAAAGC 1860  
 AAGAAATGCG ATTTTACAGC CTTCCTGAAT TGAAGAGTG GAAGAGTTCT ACTCCAATC 1920  
 ATAAAAAATG GAAAGTCAAA TATTACAAG GTTGGGCAC CAGCACATCA AAGGAGCTA 1980  
 AAGAATACTT TGCAGATATG AAAAGACATC GTATCCAGTT CAAATATTCT GGTCTGAAG 2040  
 ATGATGCTGC TATCAGCCTG GCCTTTAGCA AAAACAGAT AGATGATCGA AAGGAATGGT 2100  
 TAACTAATTT CATGGAGGAT AGAAGACAA GAAAGTTACT TGGGCTTCT GAGGATTACT 2160  
 TGTATGGACA AACTACCACA TATCTGACAT ATAATGACTT CATCAACAAG GAACTTATCT 2220  
 TGTCTCTAAA TTCTGATAAC GAGAGATCTA TCCCTCTAT GGTGATGGT TTGAAACCAG 2280  
 GTCAGAGAAA GGTTTTGTGT ACTTGCTTCA AACGGAATGA CAAGCGAGAA GTAAAGGTTG 2340  
 CCCAATTAGC TGGATCAGTG GCTGAAATGT CTCTTATCA TCATGGTGAG ATGTCACTAA 2400  
 TGATGACCAT TATCAATTG GCTCAGAATT TTGTGGGTAG CAATAATCTA AACCTCTTC 2460  
 AGCCATTGG TCAATTGCT ACCAGGCTAC ATGGTGGCAA GGATTTCTGT AGTCCAAGAT 2520  
 ACATCTTTAC AATGCTCAGC TCTTGGCTC GATTGTTATT TCCACCAAAA GATGATCACA 2580  
 CGTTGAAGTT TTATATGAT GACAACGAGC GTGTTGAGCC TGAATGTTAC ATTCTTATTA 2640  
 TTCCCATGGT GCTGATAAAT GGTGCTGAAG GAATCGGTAC TGGGTGGTCC TGCAAAATCC 2700  
 CCAACTTGA GCTTCCAGT TACAAGAACT TCAAGGTGAT TATTGAAGAA CTGGCTCCAA 2760  
 CTTTGCATAT GATTAGTGGT GAAGTAGCTA TTCTTAATTC TACAACCAT GAAATCTCAG 2820  
 ATCAATATGT CAGAACATGG ACCCAGACAT ACAAGAACAA AGTTCTAGAA CCCATGTTGA 2880  
 AGCTTCCCGT AATGCTCAGC CCTCTCATAA CAGACTATAG GGAATACCAT ACAGATACCA 2940  
 ATGGCACCGA GAAGACACCT CCTCTCATAA CAGACTATAG GGAATACCAT ACAGATACCA 3000  
 CTGTGAAATT TGTGTGAAAG ATGACTGAAG AAAAAGTGG AGAGGCAGAG AGAGTTGGAC 3060  
 TACACAAAGT CTTCAAACTC CAAACTAGTC TCACATGCAA CTCTATGGTG CTTTGTGACC 3120  
 ACGTAGGCTG TTTAAAGAAA TATGACACGG TGTGGATAT TCTAAGAGAC TTTTGTGACC 3180  
 TCAGACTTAA ATATTATGGA TTAAGAAAAG AATGGCTCCT AGGAATGCTT GGTGCTGAAT 3240  
 CTGCTAAACT GAATAATCAG GCTCGCTTTA TCTTAGAGAA AATAGATGGC AAAATATCA 3300

5	10	15	20	25	30	35	40
TTGAAATAA	GCCTAAGAAA	GAATTAATTA	AAGTTCGTGAT	TCAGAGGGGA	TATGATTCCG	3360	
ATCCTGTGAA	GGCTGTGAAA	GAAGCCGAGC	AAAAGGTTCC	AGATGAAGAA	GAAATGAAG	3420	
AGAGTGACAA	CGAAAAGGAA	ACTGAAAAGA	GTGACTCCGT	AACAGATTCT	GGACCAACCT	3480	
TCAACTATCT	TCTTGATATG	CCCTTCTGGT	ATTAAACCAA	GGAAAAGAAA	GATGAACCT	3540	
GCAGGCTAAG	AAATGAAAAA	GAACAAGAGC	TGGACACATT	AAAAAGAAAG	AGTCCATCAG	3600	
ATTTGTGGAA	AGAAGACTTG	GCTACATTTA	TTGAAGAATT	GGAGGCTGTT	GAAGCCCAAG	3660	
AAAAACAAGA	TGAACAAGTC	GGACTTCCTG	GGAAAGGGGG	GAAGGCCAAG	GGGAAAAAAA	3720	
CACAAATGCG	TGAAGTTTGG	CCTTCTCCGC	GTGGTCAAGG	AGTCATTCCA	CGAATAACCA	3780	
TAGAAATGAA	AGCAGAGGCA	GAAAAGAAAA	ATAAAAAGAA	AATTAAGAA	GAAAATACCTG	3840	
AAGGAAGCCC	TCAAGAGAGT	GGTGTGGAAC	TAGAAGGCCT	AAAAACAAGA	TTAGAAAAGA	3900	
AACAGAAAAG	AGAACCAGGT	ACAAAGACAA	AGAAACAAC	TACATTGGCA	TTTAAGCCAA	3960	
TCAAAAAGG	AAAGAAGAGA	AATCCCTGGC	CTGATTGAGA	ATCAGATAGG	AGCAGTGACG	4020	
AAAGTAATTT	TGATGTCCCT	CCACGAGAAA	CAGAGCCAGC	GAGAGCAGCA	ACAAAAACAA	4080	
AATTCACAAT	GTATTGGAT	TCAGATGAAG	ATTTCTCAGA	TTTGTATGAA	AAAACTGATG	4140	
ATGAAGATT	TGTCCCATCA	GATGCTAGTC	CACCTAAGAC	CAAACTCTCC	CCAAAACCTTA	4200	
GTAAACAAGA	ACTGAAACCA	CAGAAAAGTG	TGCTGTGAGA	CCTTGAAGCT	GATGATGTTA	4260	
AGGGCAGTGT	ACCACTGTCT	TCAAGCCCTC	CTGCTACACA	TTTCCCAGAT	GAAACTGAAA	4320	
TTACAAACCC	AGTTCCTAAA	AAGAATGTGA	CAGTGAAGAA	GACAGCAGCA	AAAAGTCAAGT	4380	
CTTCCACCTC	CACATCCGGT	GCCAAAAAAA	GGGCTGCCCC	AAAAGGAAC	AAAAGGGATC	4440	
CAGCTTTGAA	TTCTGGTGTG	TCTCAAAAGC	CTGATCCTGC	CAAAACCAAG	AATCGCCGCA	4500	
AAAGGAAGCT	ATCCACTCTA	GATGATTCTG	ACTCTAATTT	TGAGAAAAAT	GTTTCGAAAG	4560	
CAGTCACAAG	CAAGAAATCC	AAGGGGAGAG	GTGATGACTT	CCATATGGAC	TTTGACTCAG	4620	
CTGTGGCTCC	TCGGGCAAAA	TCTGTACGGG	CAAGAAACCC	TATAAAGTAC	CTGGAAGAGT	4680	
CAGATGAAGA	TGATCTGTTT	TAAAATGTGA	GGGATTTATT	TTAAGTAATT	ATCTTACCAA	4740	
GGCCAAAGCT	GGTTTAAAG	TTACCTGAAG	CTCTTAACTT	CCTCCCTCT	GAATTTAGTT	4800	
TGGGAAGGT	GTCTTTAGTA	CAAGACATCA	AAGTGAAGTA	AAGCCCAAGT	GTTCTTTAGC	4860	
TTTTTATAAT	ACTGCTCTAA	TAGTGACCAT	CTCATGGGCA	TTGTTTCTTT	CTCTGCTTTG	4920	
TCGTGTTTT	GAGTCTGCTT	TCTTTGTCT	TTAAAACCTG	ATTTTAAAGT	TCTTCTGAAC	4980	
TGTAGAAATA	GCTATCTGAT	CACCTCAGCG	TAAAGCAGTG	TGTTTATTAA	CCATCCACTA	5040	
AGCTAAACT	AGAGCAGTTT	GATTAAAG	TGTCACCTCT	CCTCCTTTTC	TACTTTCACT	5100	
AGATATGAGA	TAGAGCATAA	TTATCTGTTT	TATCTTAGTT	TTATACATAA	TTTACCATCA	5160	
GATAGAACTT	TATGTTCTTA	GTACAGATAC	TCTACTACAC	TCAGCCTCTT	ATGTGCCAAG	5220	
TTTTTCTTTA	AGCAAGGAGA	AATTGCTCAT	GTCTCTCATC	TTCTCAATC	ATCAGAGGCC	5280	
AAAGAAAAAC	ACTTTGGCTG	TGCTATAAC	TTGACACAGT	CAATAGAATG	AAGAAAATTA	5340	
GAGTAGTTAT	GTGATTATTT	CAGCTCTTGA	CCTGTCCCTC	CTGGCTGCCT	CTGAGTCTGA	5400	
ATCTCCCAA	AGAGAAACCC	AATTCTAAG	AGGACTGGAT	TGCAGAAAGC	TCGGGGACAA	5460	
CATTGTATCC	AAGATCTTAA	ATGTTATATT	GATAACCATG	CTCAGCAATG	AGCTATTAGA	5520	
TTTATTGTTG	GAAATCTCCA	TAATTTCAAT	TTGTAAACTT	TGTTAAGACC	TGCTACATTT	5580	
GTATATGTTG	TGTGACTTGA	GTAATGTTAT	CAACGTTTTT	GTAATATATT	ACTATGTTTT	5640	
TCATTAGCT	AAATCCCAAC	AATTTGTGAC	TTTAATAAAA	TGTTCTAAAC	ATTGC		

Seq ID NO: 178 Protein sequence:  
Protein Accession #: NP\_001058.1

45	1	11	21	31	41	51	
	MEVSPLOPVN	ENMQVNKIKK	NEDAKRLLSV	ERIYQKKTQL	EHILLRPDPT	IGSVELVTQQ	60
	MWYVEDVGI	NYREVTFVPG	LYKIFDEILV	NAADNKQRPD	KMSCIRVTID	PENNLISWIN	120
	NGKGIPVVEH	KVEKMYVPAL	IFGQLLTSSN	YDDDEKQVVG	GRNGYGAKLK	NIFSTKFTVE	180
50	TASREYKMF	KQTMWDMNGR	AGEMELKPFN	GEDYTCITFQ	FDLSKFKMQS	LDRDIALVMV	240
	RRAYDIAGST	KDVKVFLANG	KLPVKGFRSY	VDMLKDKLID	ETGNSLKVH	EQVNERNEVC	300
	LTMSEKGPQ	ISFVNSIATS	KGGREVDVVA	DQIVTKLVDD	VKKKNKGGVA	VKAHQVKNHM	360
	WIFVNALINE	PTFDSQTKEN	MTLQPKSFGS	TCQLSEKFIK	AAIGCGIVES	ILNWKFKKQ	420
	VQLNKKCSAV	KHNRIKGIK	LDDANDAGGR	NSTECTLILT	EGDSAKTLAV	SGLGVVGRDK	480
55	YGVPFLRGKI	LNVREASHKQ	IMENASINNI	IKIVGLQYKK	NYEDEDLSKT	LRYGKIMIMT	540
	DQDQDGSNIK	GLLINFHN	WPSLLRHRFL	EEFITPIVKV	SKNKQEMAFY	SLPEFEWKS	600
	STPNHKKWV	KYKGLGTST	SKEAKYFPAD	MKRHRIFQFY	SGPEDDAIS	LAFSKKQIDD	660
	RKEWLTNFM	DRRQRLKGL	PEDYLYGQTT	TYLTYNDFIN	KELILFNSND	NERSIPSMVD	720
	GLKPGGRKVL	FTCFKRNDR	EVKVAQLAGS	VAEMSSYHHG	EMSLMNTIIN	LAQNFVGSNN	780
60	LNLLPIQIQF	GTARLHGKDS	ASPRYIFTML	SSLARLLFPP	KDDHTLKLFLY	DDNQREVEPN	840
	YIPIIFMVL	NGAEGIGTGW	SKIPNFDVR	EIVNNIRRLM	DGEEPLPMLP	SYKNFKGTIE	900
	ELAPNQYVIS	GEVAILNSTT	IEISELFPVT	WTQTYKEQVL	EPMLNGTEKT	PPLITDYREY	960
	HTDTTVKFFV	KMTBEKLAEA	ERVGLHKVFK	LQTSLTCSNM	VLFHDVGLCK	KYDVTDLILR	1020
	DFPELRLLKY	GLRKEWLLGM	LGAESAKLNN	QARFILEKID	GKIIENKPK	KELIKVLIQR	1080
65	GVDSDPVKAW	KEAQKQVPE	EENESDNEK	ETEKSDSVTD	SGPTFNLYLD	MPLWYLTKEK	1140
	KDELCLRLNE	KEQELDTLKR	KSPSDLWKED	LATFIELEEA	VEAKKQDEQ	VGLPGKGGKA	1200
	KGKTKMAEV	LPSPRQQRVI	PRITIEKAE	AEKKNKKIK	NENTEGSPQE	DGVELEGLKQ	1260
	RLEKKQKREP	GKTKTKQTTL	AFKPIKKGKK	RNPWPDSESD	RSSDESNDPV	PPRETEPRRA	1320
	ATKTKFTMDL	DSDEDFSPFD	EKTDDDFVP	SDASPPKTKT	SPKLSNKKEL	PQKSVVSDLE	1380
70	ADDVKGSVPL	SSSPPATHPF	DETEITNFPV	KKNVTVKKTA	AKSQSSTSTT	GAKKRAAPFG	1440
	TKRDPALNSG	VSQKPDPAKT	KNRRKRKPS	SDSDSNFEK	IVSKAVTSKK	SKGESDDPHM	1500
	DFDSAVAPRA	KSVRAKKPIK	YLEESDEDDL	F			

Seq ID NO: 179 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-7095

80	1	11	21	31	41	51	
	CACACATACG	CAGCAGCAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGAGGGG	CGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCATT	180
	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
85	CTTGTGAAAG	AGATTGGCTG	GTCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420

	AACACATTCA	TTCCATAACAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTGAGGGGAG	GAGTTTCAGA	AATGGTGT	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTACTGTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTCTCACTT	600
5	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTITTTGGGA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCTGTCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
10	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTITTTTGTG	AAGTTCCTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTCAGAAC	CGAAAAATGT	TCAAGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTCAC	1140
	TGGGAAAGAC	CTCGATCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAG	1200
15	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTG	TCATAATTTT	GCTACCCAAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCTCGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACCTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGTCTACA	1500
20	AAACAAATCA	GGAAAGAGGA	ACCCAGATT	TCATCCACAA	CACACTACAA	TCGCATAGGG	1560
	ACGAAATACA	ATGAGGCCAA	GACTAACCGA	TCGCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTATG	TTCCCAATTG	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCTCT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGTG	1740
	GAGGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACCTG	TTCTTAGATC	TCCATATATG	1800
25	AACTTGTGGG	GGACTGCAGA	ATCCTTAAAT	ACAGTTTCTA	TAACAGAATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	GCTTCCCTAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAGAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGTAGT	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTCAGAAGAA	TCACATAAAG	ATCCTTCTAT	GGAGGGAAAT	2100
30	GTGTGTTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCC	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACCTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTTCTG	CAGGCCCATG	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTGGCTTA	CTTCCCAACT	GAGGTAAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTTGGT	CTCCACGGTC	AACGTGGTAT	ACTGCGAGAC	AACCCACCG	2400
35	GTATACAATG	GTGAGACACC	TCTTCAACCT	TCTTACAGTA	GTGAAGTCTT	TCTCTAGTCT	2460
	ACCCCTTTGT	TGCTTGACAA	TCAGATCTCT	AACTACTACC	CTGCTGCTTC	AAGTAGTGAT	2520
	TGCGCTTGGC	ATGCTGACCC	TGTATTTCCC	AGTGTGATG	TGCTATTGTA	ATCCATCCCTG	2580
	TCTTCTCATG	ATGGTGACCC	TTTGCTTCCA	TTTCTCTCTG	CTTCTCTCAG	TAGTGAATTG	2640
	TTTCCGCCATC	TGCAATACAG	TTCTCAATATC	CTTCCACAA	TTACTTCAGC	TACCGAGAGT	2700
40	GATAAGGTGC	CCTTGCATGC	TTCTCTGCCA	GTGGCTGGGG	GTGATTTGCT	ATTAGAGCCC	2760
	AGCCTTGGCTC	AGTATTTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGGTAGTG	CTTCTGGTGT	TCTTTATAAA	AGCCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTAGTG	CCATGATGCA	TGCAAGTTCT	TCAGGGCCCTG	AACTTCTTCA	TGCTTGTCT	2940
	GATAATGAGG	GCTTCCCAACA	CATCTTCACT	GTTCCTTACA	GTTCGTCAAT	ACCTGTGAT	3000
45	GATTCTGTGG	GTGTAACCTTA	TCAGGGTCTC	TTATTTAGCG	GGCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGAATGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	TGGCTGACAG	CCCTTAAACAT	TTCTTCACT	GTTCCTGTAG	CTGAATTTAC	ATATACAACA	3240
	CTGTGTGTTG	GTGATGATAA	TAAGGGGCTT	TCTAAAAGTG	AAATAATATA	TGGAAATGAG	3300
50	ACTGAACTGC	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAGG	CACAGTCAATG	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGGCTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGGCAT	GTITTCAGGG	TCCCTTGTCT	ATACCAACCAC	TAAAGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACTTTTCA	TTCAACCTAC	ACATACTGTC	3540
	TCTCAAGCAT	TGCTGACATC	TTGCTTAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
55	TCTCTGACCC	CTGCTTCTAG	TGAAATGTTA	TCTCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATG	CTACAACTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAAACCTG	TCTTCCAGCT	GTGCCAGTG	ATCCAATATT	GGTTGAAACC	3780
	CCCAAAAGTG	ATAAAATAG	TTTCAATAG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
	AGTGAAGACA	TGCTGCATCT	TACATCTGTA	CCAGTTTGTG	ATGTGTGCGC	TACTTCTCAT	3900
60	ATGCATCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTITTTGTTAA	AAAGTGAAG	TTCCCAACCA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TGTGTTCCAA	CGGCCAATTT	GGAGATTAA	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CGTTTATATC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAAGCTTATA	4140
	CATTTCGATG	AAATTTTAA	CTCCACCAAA	AGTTCTGTGA	CTGGTAAGGT	ATTGCTGGT	4200
65	ATTCCACAG	TTGCTTCTGA	TACATTGTA	TCTACTGATC	ATTCGTGCTC	TATAGGAAAT	4260
	GGGCAATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCGTAAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
	AGAGGTAGTG	ATGGCTTATC	CATTATAAG	TGTATGTCAT	GCTCATCTTA	TAGAGAATCA	4500
70	CAGGAAAAGG	TAATGAATGA	TTGACACACC	CACGAAAACA	GTCTTATGGA	TCAGAAATAAT	4560
	CCAATCTCAT	ACTCACTATC	TGAGAACTCT	GAAGAAAGATA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAAATGATG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCGACAA	GTGATGAAGA	AAGTGGATCA	4800
75	GGGCAAGGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTGTGA	4860
	GACACTAATG	AAAAGATGTC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AATAACTCCT	4920
	GGATTCCAC	AGTCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAGT	GTTCACGTT	4980
	TCAGAGGCGAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	5040
	GAATCCGAGA	AGAAGGCACT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
80	CTAGTGGTTC	TTTGGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAAACCC	TATCTTTCCA	5220
	ATTTCAAGATG	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAAATT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	5340
	CAGAGCTGTA	CTGTTGACTT	AGGTATTACA	GCAGACAGCT	CCACCAACCC	AGACAAACAG	5400
	CACAGAAATC	GATACATAAA	TATCGTTGCC	TATGATCTATA	GCAGGGTTAA	GCTAGCACAG	5460
85	CTTGCTGAAA	ACTGAGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	5520
	AACAGACCAA	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	5580
	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTATTGTGCA	TGATAACAAA	CCTCGTGGAG	5640



	AAAGGAAGGA	GAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	5700
	TTTCCTGGTC	CTCAGAGAG	TGTGCAAGTG	CTTGCCCTAT	ATACTGTGAG	GAATTTTACT	5760
5	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	5820
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	5880
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTGTGCGTC	5940
	CAGTGCAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	6000
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAACACAT	CCGTTCACAA	6060
	AGAAATTATT	TGGTACAAAC	TGAGGAGCAA	TATGCTTTCA	TTCATGATAC	ACTGGTTGAG	6120
10	GCCATACCTA	GTAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTCATGCCCTA	TGTTAATGCA	6180
	CTCCTCATTC	CTGGACGAGC	AGGCAAAACA	AAGCTAGAGA	AACAATTCCA	GCTCCTGAGC	6240
	CAGTCAAAAT	TACAGCAGAG	TGACTATTCT	GCAGCCCTAA	AGCAATGCCA	CAGGGAAGAG	6300
	AATGGAACCT	CTTCTATCAT	CCCTGTGGAA	AGATCAAGGG	TTGGCATTTC	ATCCCTGAGT	6360
	GGAGAAGGCA	CAGACTACAT	CAATGCCCTCC	TATATCATGG	GCTATTACCA	GAGCAATGAA	6420
	TTATCATATTA	CCAGCACCCC	TCTCCTTCAT	ACCATCAAGG	ATTCTGGGAG	GATGATATGG	6480
15	GACCATAAAT	CCCAACTGTG	GGTTATGATT	CCTGATGGCC	AAACATGGCC	AGAAGATGAA	6540
	TTTGTTTACT	GGCCAAATAA	AGATGAGCCT	ATAAATTGTG	AGAGCTTTAA	GGTCACTCTT	6600
	ATGGCTGAAG	AACACAAATG	TCTATCTAAT	GAGGAAAAAC	TTATAATTCA	GGACTTTATC	6660
	TTAGAAGCTA	CACAGGATGA	TTATGTACTT	GAAGTGAGGC	ACTTTCAGTG	TCCTAAATGG	6720
20	CCAAATCCAG	ATAGCCCCAT	TAGTAAACT	TTTGAACCTA	TAACTGTTAT	AAAAGAAGAA	6780
	GCTGCCAATA	GGGATGGGCC	TATGATTGTT	CATGATGAGC	ATGAGGAGT	GACGGCAGGA	6840
	ACTTTCTGTG	CTCTGACAA	CCTTATGCAC	CAACTAGAAA	AAGAAATTC	CGTGGATGTT	6900
	TACCAAGTAG	CCAAGTAGAT	CAATCTGATG	AGGCCAGGAG	TCCTTGTCTA	CATTGAGCAG	6960
	TATCAGTTTC	TCTACAAAGT	GATCCTCAGC	CTTGTGAGCA	CAAGGCAGGA	AGAGAAATCA	7020
25	TCCACCTCTC	TGGACAGTAA	TGGTGCAGCA	TTGCCTGATG	GAAATATAGC	TGAGAGCTTA	7080
	GAGTCTTTAG	TTTAACACAG	AAAGGGGTGG	GGGGACTCAC	ATCTGAGCAT	TGTTTTCTTC	7140
	TTCTAAAT	TAGGCAGGAA	AATCAGTCTA	GTTCTGTATT	CTGTGTATT	CCCATCACT	7200
	GACAGTAATC	TTATGACAT	AGGATTCTGC	CGCCAAATTT	ATATCATTA	CAATGTGTGC	7260
	CTTTTGTCAA	GACTTGTAA	TTACTTATTA	TGTTTGAAT	AAATGATTG	AATTTTACAG	7320
30	TATTTCTAAG	AATGGAATG	TGGTATTTTT	TTCTGTATTG	ATTTTAACAG	AAAATTTCAA	7380
	TTTATAGAGG	TTAGGAATTC	CAAACTACAG	AAAATGTTTG	TTTTTAGTGT	CAAAATTTTA	7440
	GCTGTATTTG	TAGCAATTAT	CAGGTTTGCT	AGAAATATAA	CTTTTAATAC	AGTAGCCTGT	7500
	AAATAAACA	CTCTTCCATA	TGATATTCAA	CATTTTACAA	CTGCAGTATT	CACCTAAAGT	7560
	AGAAATAATC	TGTTACTTAT	TGTAATACT	GCCCTAGTGT	CTCCATGGAG	CAAAATTTATA	7620
35	TTTATAATG	TAGATTTTTA	TATTTTACTA	CTGAGTCAAG	TTTTCTAGIT	CTGTGTAATT	7680
	GTTTAGTTTA	TAGACGTAGT	TCATTAGCTG	GTCCTTACTCT	ACCACTTTTC	TGACATTGTA	7740
	TTGTGTTACC	TAAGTCATTA	ACTTTGTTTC	AGCATGTAAT	TTTAACTTTT	GTGGAATAA	7800
	GAAATACCTT	CATTTTGAAA	GAGTTTTTTA	TGAGATAAC	ACCTTACCAA	ACATTGTTCA	7860
40	AATGGTTTTT	ATCCAAGGAA	TTGCAAAAAT	AAATATAAAT	ATTGCCATTA	AAAAAATAA	7920
	AAAAAATAA	AAAAAATAA	AAAA				

Seq ID NO: 180 Protein sequence:  
Protein Accession #: Bos sequence

45	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRID	WANGYYRQQR	KLVEEIGWSY	TGALNQKNG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKRL	KPQGWDRKSL	ENTFIHNTGK	TVELINTNDY	RVSGGVSEMV	120
	FKASKITFW	GKCNMSDDGS	EHSLEGQKFP	LEMQIYCFDA	DRFSSPERAV	KGKGLRLALS	180
50	ILFVETGTEEN	LDPKAIIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSDKYIY	YNGSLTSPPC	240
	TDVDWIVFK	DTVISSESOL	AVFCEVLIMQ	QSGYVLMMDY	LQNNFREQQY	KFSRQVFPSSY	300
	TGKEEIEHAV	CSSEPEVQOA	DPENYTSLLV	TWERPRVVVD	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAALLNLPL	NMSYVLQIVA	ICTNGLYGYK	SDQLIVDMPT	DNPELDLPE	420
	LIGTEEIIKE	EEEGKDIIEG	AIVNPGRDSA	TNQIRKCBPO	ISTTHYNNRI	GTRYNEAKTN	480
55	RSPTRGSEPS	GKGDVPNTSL	NSTSQPVTKL	ATEKDISLTS	QTVELEPHPT	VEGTSASLND	540
	GSKTVLRSPH	MNLGSGTASL	NTVSIITEYE	ESLLTSFKLD	TGAEDSGSGS	PATSAIPFIS	600
	ENISQGYIFS	SENPEITITD	VLIPESARNA	SEDSTSSGSE	ESLKDPSEMG	NVWFPSSTDI	660
	TAQPDVGSGR	ESFLQNTYTE	IRVDESEKTT	KSPSAGPVMS	QGPVTDLEM	PHYSTPAYFP	720
	TEVTPHAFPT	SSRQQLVST	VNVVYSQTTQ	PVYNGETPLQ	PSYSSVFPPL	VTPLLLDNQI	780
60	LNTTPHASSS	DSALHATPFF	PSVDVSFESI	LSSYDGAPLL	PFSSASFSSE	LFRHLHTVSQ	840
	ILPQVTSATE	SKKVPFHASL	PVAGDILLLE	PSLAQYSDVL	STTHAASETL	EFGSESGVLVY	900
	KTLMPSQVEP	PSGDAMMHAR	SSGPEPSYAL	SDNEGSOHIF	TVSYSSAIFV	HDSVGVTVYQ	960
	SLPSGSPSHIP	IPKSSLITPT	ASLQPTHAL	SGDGEWSGAS	SDSEFLLPDT	DGLTALNISS	1020
	PVSVAEFTYT	TSVFGDDNKA	LSKSEIIYGN	ETELQIPSPN	EMVYPSESTV	MPNMYDNVNM	1080
65	LNASLQETSV	SISSTKGMFP	GSLAHTTKV	FDHEISQVPE	MNPSVQPTET	VQSAGDTSIL	1140
	KPVLANSSEP	ASSDPASSEM	LSPSTQLLFY	ETSASFSTEV	LLQPSFQASD	VDTLKTVLP	1200
	AVPSDPIIVE	TPKVDKISST	MLHLIVNSA	SSENMLHSTS	VPVFDVSPTS	HMSASLQGL	1260
	TISYASEKEY	PVLLKSESSH	QVPSLYSND	ELPQTANLEI	NQARPPKGRH	VFATPULSID	1320
	EPLMTLINKL	IHSDEILTST	KSSVTGKVFA	GIFTVASDTF	VSTDHSPVIG	NGHVAITAVS	1380
70	PHRDGVSST	KLLFPKATS	ELSHSAKSDA	GLVGGGEDGD	TDDGDGDDDD	DRGSDGLSIH	1440
	KCMSCSSYRE	SQEKVMNDSD	THENSLMDQN	NPISYLSSEN	SEEDNRVTSV	SSDSQTGMDR	1500
	SPGKSPSANG	LSQKHNDGKE	ENDIQTSAL	LPLSPESKAW	AVLTSDEESG	SGQGTSDSLN	1560
	ENETSTDFSF	ADTNEKDADG	ILAGDSEIT	PGFPQSPSTSS	VTSENSEVFH	VSEAEASNSS	1620
	HESRIGLABG	LESEKKAIVP	LVIVSALTFI	CLVVLGILY	YWRKCPQTAH	FYLEDSTSPR	1680
75	VISTPPTPIF	PISDDVGAIP	IKHFPKHVAD	LHASSGFTEE	FETLKEFYQE	VQSCTVDLGI	1740
	TADSSNHPDN	KHKRYINIV	AYDHSRVKLA	QLAEKDGKLT	DYINANYVDG	YNRPKAYIAA	1800
	QGPKLSTAE	FWRMIWEENV	EVIVMITNLV	EKGRKCDQY	WPADGSEYEG	NFLVTKQSVQ	1860
	VLAYITVRNF	TLRNTKIKKG	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTFVRKAAY	1920
	AKRHAVGPVV	VHCSAGVGRT	GTIVLDSML	QQIQHSTVNM	IFGFLKHIRS	QRNYLVQTEE	1980
80	QYVFIHDTLV	BAILSKETEV	LDSHIRAYVN	ALLIPGPAGK	TKLEKQFQLL	SQSNIQQSDY	2040
	SAALKQCNR	KNRTSSIIIP	ERSRVGISSL	SGESTDYINA	SYIMGYQQSN	EFITQHPILL	2100
	HTIKDFWRMI	WDHNAQLVVM	IPDQONMAED	EFVYWPKNDE	PINCESFKVT	LMAEHEKCLS	2160
	NEEKLIIDPF	ILEATQDDYV	LEVHRFQCPK	WPNPDSPISK	TFELISVIKE	BAANRDGPMI	2220
	VHDSHGGVTA	GTFCALTTLM	HQLEKENSVD	VYQVAKMINL	MRPGVFADIE	QYQFLYKVL	2280
85	SLVSTRQREN	PSTSLDSNGA	ALPDGNIAES	LESIV			

Seq ID NO: 181 DNA sequence  
Nucleic Acid Accession #: Bos sequence



	1	11	21	31	41	51	
5	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGCGGAGGGG	CGCGAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGC CG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
10	CTTGTGAAG	AGATTGGCTG	GTCTATACAC	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
	AACACATTCA	TTCATAAACAC	TGGGAAAAAC	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTCCAGCGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
15	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAGGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAAG	TTGGGACAGA	AGAAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTGTAGT	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCGT	CCAACTCAA	CTGACAAGTA	TTACATTTAC	840
20	AATGSCATAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTCTTTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCACTA	AGCAGTTTGT	1080
	AGTTTCAGAA	TCATGCTGAT	TGAGGCTGAC	CCAGAGAAAT	ATACCAGCCT	TCTTGTATCA	1140
25	TGGGAAAGAC	CTGAGTCTGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTG	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAAATACAG	GACCAACTGA	TTGTTCAGAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAAGTG	AAGAAATAAT	CAAGGAGGAG	1440
30	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AGGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACAGTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACGCCCACT	TCACACTGTG	1740
35	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAAGTG	TTCTTAGATC	TCCACATATG	1800
	AACCTGTGCG	GAGTTCGAGA	ATCTTAAAT	ACAGTTTCTA	TACAGAAATA	TGAGGAGGAG	1860
	AGTTTATTTA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTTCC	1980
	GAAGAACCCAG	AGACAAATAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
40	GAAGATTCAA	CTTCACTAGG	TTCAGAGAA	TCACTAAGGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGCTTTC	CTAGCTTCTC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCTTTTCTCG	CAGGCCCACT	GATGTCCAGG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CCTTTGCCCT	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
45	TCCAGACATC	AGGATTGGGT	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
	GTATACATG	CAGAGGCCAG	TAATAGTAGC	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAAGGC	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTATC	2520
	TGTCTAGTGG	TTCTGTGGGG	TATTTCTATC	TACTGGAGGA	AATGCTTCCA	GACTGCACAC	2580
	TTTTACTTAG	AGGACAGTAG	ATCCCTTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTT	2640
50	CCAAATTCAG	ATGATGGGCG	AGCAATTCOA	ATAAAGCACT	TTCCAAGCA	TGTTGCAGAT	2700
	TTACATGCAA	GTAGTGGGTT	TACTGAAGAA	TTTGAGACAC	TGAAGAGATT	TTACAGGAA	2760
	GTGCAGAGCT	GTACTGTGTA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCGAGACAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAAGCTAGCA	2880
	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATTATATCA	ATGCCAATTA	TGTTGATGGC	2940
55	TACAACAGAC	CAAAAGCTTA	TATTGCTGCC	CAAGGCCCACT	TGAATTCAC	AGCTGAAGAT	3000
	TTCTGGAGAA	TGATATGGGA	ACATAATGTG	GAAGTTATTG	TCATGATAAC	AAACCTCTGT	3060
	GAGAAAGGAA	GGAGAAATAG	TGATCAGTAC	TGGCCTGCGG	ATGGGAGTGA	GGAGTACGGG	3120
	AACCTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCCCT	ATTATACTGT	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAAAAT	AAAAAGGGGC	TCCAGAGAA	GAAGACCCAG	TGGAGCTGTG	3240
60	GTCAACAGAT	ATCACTACAC	CGAGTGGCCT	GACATGGGAG	TACCAAGATA	CTCCCTGCCA	3300
	GTGCTGACCT	TTGTGAGAAA	GGCAGCCTAT	GCCAGGCCCT	ATGCACTGGG	GCCTGTGTC	3360
	GTCCACTGCA	GTGCTGGAGT	TGGAAGAAC	GGCACAATTA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTC	AACACGAAGG	AACGTCAAC	ATATTGGCT	TCTTAAACCA	CATCCGTTCA	3480
	CAAGAGAAAT	ATTGGGTACA	AACGTAGGAG	CAATATGCTT	TCAATCATGA	TACACTGGTT	3540
65	GAGGCCATAC	TTAGTAAAGA	AACGTAGGAG	CTGGACAGTC	ATATTCAATG	CTATGTTAAT	3600
	GCACTCTCTA	TTCTTGGACC	AGCAGGCCAA	ACAAAGCTAG	AGAAACAATT	CCAGCTCCTG	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGACTAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AAGAATCGAA	CTTCTTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTGGCAT	TTCACTCCCTG	3780
	AGTGGAGAG	GCACAGACTA	CATCAATGCC	TCCTATATCA	TGGGCTATTA	CCAGAGCAAT	3840
70	GAATTATCA	TTACCCAGCA	CCCTCTCTCT	CATACCATCA	AGGATTCTCT	GAGGATGATA	3900
	TGGGACCAT	ATGCCCAACT	GGTGGTTATG	ATTCTGTATG	GCCAAAACAT	GGCAGAAGAT	3960
	GAATTGTGTT	ACTGGCCAAA	TAAAGATGAG	CCTATATAAT	GTGAGAGCTT	TAAGGTCACT	4020
	CTTATGGCTG	AAGAACAACA	ATGTCTATCT	AATGAGGAAA	AACCTATAAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGGA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
75	TGGCCAAATC	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAGAA	4200
	GAAGCTGCCA	ATAGGGATGG	GCCTATGATT	GTTCATGATG	AGCATGGAGG	AGTGAOGGCA	4260
	GAAGCTTTCT	GTGCTCTGAC	AACCTTTATG	CACCAACTAG	AAAAAGAAAA	TTCCGTGGAT	4320
	GTTTACAGG	TAGCCAGAGT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCCTC	AGCCTTGTGA	GCACAAGGCA	GGAGAGAAAT	4440
80	CCATCCACTT	CTCTGGACAG	TAATGGTGCA	GCATTGCCCT	ATGGAATAT	AGCTGAGAGC	4500
	TTAGAGTCTT	TAGTTTAAACA	CAGAAAGGGG	TGGGGGGACT	CACATCTGAG	CATTGTTTTT	4560
	CTCTCTCTAA	AATTAGGCAG	GAATAATCAGT	CTAGTTCTGT	TATCTGTGTA	TTTCCCATCA	4620
	CCTGACAGTA	ACTTTCACTA	CATAGGATTTC	TGCCGCCAAA	TTTATATCAT	TAACAATGTG	4680
	TGCCCTTTTG	CAAGACTTGT	AATTACTTTA	TTATGTTTGA	ACTAAATGA	TTGAATTTTA	4740
85	CAGTATTTCT	AAGAATGGAA	TTGTGGTATT	TTTTTCTGTA	TTGATTTTAA	CAGAAATTTT	4800
	CAATTTATAG	AGGTTAGGAA	TTCCAAACTA	CAGAAATATG	TTGTTTTTAG	TGTCAAATTT	4860
	TTAGCTGTAT	TTGTAGCAAT	TATCAGGTTT	GCTAGAAATA	TAACTTTTAA	TACAGTAGCC	4920
	TGTAATAAAA	ACACTCTTCC	ATATGATATT	CAACATTTTA	CAACTGCACT	ATTCACCTAA	4980

AGTAGAATA ATCTGTACT TATTGTAAT ACTGCCCTAG TGTCTCCATG GACCAAATTT 5040  
 ATATTATATA TTGTAGATT TTATATTTTA CTACTGAGTC AAGTTTTCTA GTTCTGTGTA 5100  
 ATTGTTTAGT TTAATGACGT AGTTCATTAG CTGGTCTTAC TCTACCAAGT TTCTGACATT 5160  
 GTATTGTGTT ACCTAAGTCA TTAACCTTGT TTCAGCATGT AATTTTAACT TTGTGGGAAA 5220  
 ATAGAAATAC CTTCATTTTG AAAGAAGTTT TTATGAGAAT AACACCTTAC CAAACATTGT 5280  
 TCAAAATGGTT TTTATCCAAG GAATTGCAAA AATAAATATA AATATTGCCA TTAATAAAAA 5340  
 AAAAAAAAAA AAAAAAAAAA AAAAAA

Seq ID NO: 182 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MRILKRFLAC IQLLCVCRLD WANGYRQQR KLVEEIGWSY TGALEQKNGW KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKDL KFGQMDKTSI ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFHW GKCMSSSDGS EHSLEQKXP LEMQIYCPDA DRPSSFEAV KKGKLRALS 180  
 ILFEVGTZEN LDFKAIIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMO QSGYVLMMDY LQNNFREQQY KFSRQVFSY 300  
 TKKEEIEHAV CSSEFENVQA DPENYTSLLV TWERPRVVDY TMIEKPAVLV QQLDGEDQTK 360  
 HEFLIDGVQD LGAILNLLP NMSVVLQIVA ICTNGLYGRY SDQLIVDMPT DNPELDFPE 420  
 LIGTEETIKE EEEGRDIEEG AIVNFRDASA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDVFNLSL NSTSQPVTKL ATEKDISLTS QVTTELPPHT VEGTSASLND 540  
 GSKTVLRSPH MNLGTAESL NTVSITEYEE ESLITSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETITVD VLIPESARNA SEDSTSSGSE ESLKDPMSMEG NVWFPSSTDI 660  
 TAQPDVGSGR BSFLQNTNTE IRVDESEKTT KSFSAGFVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTHFAFTP SSRQODLVST VNVVYSQTTQ PVYNABASNS SHESRIGLAE GLESEKKAVI 780  
 PLVIVSALTP ICLVVLVGIL IYWRKCFQTA HFYLEDSTSP RVISTPPTPI FPSDDVGA 840  
 PIKHFFPKHVA DLHASSGTFE EFETLKEFYQ EVQSCVDLQ ITADSSNHPD NKHNRYINI 900  
 VAYDHSRVKL AQLAEKDGKL TDYINANYVD GYNRPKAYIA AQGPLKSTAE DFWRMWHEHN 960  
 VEIVVMITNL VEKGRKRCQD WNPADGSEEV GNFLVTQKSV QVLAYTVTRN FTLRNTKIKK 1020  
 GSQGRPSGR VVTQYHYTQW PDMGVPEYSL PVLTFVRKAA YAKRHAVGPV VVHCSAGVGR 1080  
 TGTIVLDSM LQIQHQBTV NIFGFLKHIR SQRNLYVQTE EQYVFIHDTL VEAILSKETE 1140  
 VLDSHIHAYV NALLIPGPAG KTKLEKQFQL LSQSNIQQSD YSAALKQCNR EKNRTSIIIP 1200  
 VERSRVGISS LSGEGTDYIN ASYIMGYQOS NEFITTQHPL LHTIKDFWRM IWDHNAQLVV 1260  
 MIPDQNMAB DEFVYWNPK EPINCESFRV TLMAEEHKCL SNEEKLIQD FILEATQDDY 1320  
 VLEVRHFQCP KWPNDSPIS KTFELISVIK EEAANRDGPM IVHDEHGGVT AGTFCALITL 1380  
 MHQLEKENSU DVYQVAKMIN LMRPGVFADI EQYQFLYKVI LSLVSTRQEE NPSTSLDSNG 1440  
 AALPDGNIAB SLESIV

Seq ID NO: 183 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4494

1 11 21 31 41 51  
 CACACATACG CACGCACGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAGAG ATTTCTCTCG CTCGCCCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGGCAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCTC CGCTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCGC CCGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTTGTGTAAG AGATTGGCTG GTCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAG 300  
 AAATATCCAA CATGTAATAG CCCAAACAAA TCTCCTATCA ATATTGATGA AGATCTTACA 360  
 CAGTAAATG TGAATCTTAA GAAACTTAAA TTTCAGGTTT GGGATAAACA ATCATGGGAA 420  
 AACACATTCA TTCTAACAC TGGGAAAACA GTGGAATTA ATCTCACTAA TGACTACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTIT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AAATGCATA TGTCTCTGTA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600  
 GAGATGCAAA TCTACTGCTT TGATGCAGAG CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAAAGGGA AGTTAAGAGC TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAATTG 720  
 GAITTCAAGC CGATTATGTA TGGAGTGGAA AGTGTGTAGT GTTTTGGGAA GCAGGCTGCT 780  
 TTAGATCCAT TTAGACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTAC 840  
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAAGTTGGCT GTTTTGTGTG AAGTCTTAC AATGCAACAA 960  
 TCTGGTTATG TCATGCTGAT GGACTACTTA CAAAACAATT TTCAGAGACA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAGGAAG AGATTCTAGA AGCAGTTTGT 1080  
 AGTTCAGAAC CAGAAAATGT TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTTGTACA 1140  
 TGGGAAGAGC CTCAGTCTGT TTATGATACC ATGATTGAGA AGTTTGCACT TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTG TCAATAATTT GCTACCCAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGAGCAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCTGGAATTA ATTGGAACTG AAGAAATAAT CAAGGAGGAG 1440  
 GAAGAGGGAA AAGACATTGA AGAAGCGCT ATTGTGAATC CTGTTAGAGA CAGTGCTACA 1500  
 AACCATAACA GGAAGAGGA ACCCCAGATT TCTACCACAA CACTACTCAA TCGCATAGGG 1560  
 ACGAAATACA ATGAAGCCAA GACTAACCGA TCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620  
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCCTT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGGTACTT CAGCCTCTTT AAATGATGGC TCTAAAACCTG TTCTTAGATC TCCCATATG 1800  
 AACTTGTGCG GGAATGACGA ATCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATGGA CAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920  
 GCAACTCTG CTATCCCAAT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTTCC 1980  
 GAAACCCAG AAGCAATAAC ATATGATGTC CTATACCAG AACTCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTCATCAGG TTCAGAGAA TCACTAAAGG ATCCTTCTAT GGAGGGAAT 2100  
 GTGTGCTTCT CTAGCTTCTC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTCTTCC AGACTAATTA CACTGAGATA CGTGTGTATG AATCTGAGAA GACAACCAAG 2220  
 TCCCTTCTG CAGGCCACTG GATGTCACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280  
 CATATTCTTA CCTTTGCCTA CTTCACCACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340  
 TCCAGACAAC AGGATTGTGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCACCG 2400  
 GTATACAATG AGGCCAGTAA TAGTAGCCAT GAGTCTCGTA TTGGTCTAGC TGAGGGGTG 2460

	GAATCGGAGA	AGAAGGAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCACACCC	TATCTTTCCA	2640
5	ATTTAGATG	ATGTGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAAATT	GAGGAAAGTC	AGAGCTGTAC	TGTTGACTTA	2760
	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAAACAGC	ACAAGAAATG	ATACATAAAT	2820
	ATCGTTGCCT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCAAA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
10	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTTC	GGAGAAATGAT	ATGGGAACAT	3000
	AATGTGGAAG	TTATTGTCTAT	GATAACAAAC	CTCGTGAGAG	AAGGAAGGAG	AAATATGTAT	3060
	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAGT	TTCTGGTCTAC	TCAGAAAGAT	3120
	GTGCAAGTGC	TTGCCCTATTA	TACTGTGAGG	AAATTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGTGCA	CACAGTATCA	CTACACGCAG	3240
	TGGCTTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
15	GCCTATGCCA	AGAGCCATGC	AGTGGGGCCT	GTGTCTGTCC	ACTGCAAGTC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAAGT	3420
	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTCAACAA	GAAATTTATT	GGTACAAACT	3480
	GAGGAGCAAT	ATGCTTCTAT	TCATGATACA	CTGGTTGAGG	CCATCTTAGT	TAAAGAAACT	3540
20	GAGGTGCTGG	ACAGTCAATAT	TCATGCTTAT	GTTAATGCAC	TCCTCATTCC	TGGACGAGCA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	AGTCAAAATAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAGGAA	ATCGAACTTC	TTCTATCATC	3720
	CCTGTGGAAA	GATCAAGGGT	TGGCAITTTCA	TCCTGTAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCCCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCATTAC	CCAGCACCTT	3840
25	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATAATGC	CCAATGCTG	3900
	GTTATGATTG	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTTGTA	GAGCTTTAAG	GTCACTCTTA	TGGCTGAAGA	ACACAAATGT	4020
	CTATCTAATG	AGGAAAACT	TATAATTGAG	GACTTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	CAAAATCCAG	TAGCCCTATT	4140
30	AGTAAACTTT	TTGAACCTTAT	AAGTGTATTA	AAAGAAGGAG	CTGCCAATAG	GGATGGGCCT	4200
	ATGATTTGTT	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTTCGTGTC	TCTGACAAAC	4260
	CTTATGCACC	AATCAGAAA	AGAAAAATCC	GTGGATGTTT	ACCAAGTAGC	CAAGATGATC	4320
	AATCTGATGA	GGCCAGGAGT	CTTGTCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACTCTCT	GGACAGTAAT	4440
35	GGTGACGAT	TGCGTATGG	AAATATAGCT	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTCTCTCT	TCCTAAAATT	AGGCAGGAAA	4560
	ATCAGCTAG	TCTGTATTTC	TGTTGATTTC	CCATCACCTG	ACAGTAACTT	TCATGACATA	4620
	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	TTTTTGCAAG	ACTTGTAAAT	4680
40	TACTTATTAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
	GGTATTTTTT	TCTGTATTGA	TTTTAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTC	4800
	AAACTACAGA	AAATGTTTGT	TTTTAGTGTC	AAATTTTAG	CTGTATTTGT	AGCAATTATC	4860
	AGGTTTGCTA	GAAATATAAC	TTTTAATACA	GTAGCCTGTA	AAATAAACAC	TCTTCCATAT	4920
	GATATTCAAC	ATTTTCAAC	TGCAATATT	ACCTAAAGTA	GAAATAATCT	GTTACTTATT	4980
45	GTAAATACTG	CCCTAGTGT	TCCATGGACC	AAATTTATAT	TTATAATTGT	AGATTTTAT	5040
	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAATTG	TTTAGTTTAA	TGACGTAGTT	5100
	CATTAGCTGG	TCTTACTCTA	CCAGTTTCT	GACATTGTAT	TGTGTACCT	AAGTCATTAA	5160
	CTTGTGTTCA	GCAATGTAAT	TTAACTTTTG	TGGAAAATAG	AAATACCTTC	ATTTTGAAG	5220
	AAGTTTTTAT	GAGAAATAAC	CCTTACCAAA	CATTGTTCAA	ATGGTTTTTA	TCCAAGGAAT	5280
50	TGCAAAAATA	AATATAAATA	TTGCCATTAA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	5340
	AAA						

Seq ID NO: 184 Protein sequence:  
Protein Accession #: EOS sequence

	1	11	21	31	41	51	
55	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNBPK	60
	QSPINIDEDL	TQVNVNKKKL	KPGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSGGVSEMV	120
	FKASKITPHW	GKCNMSSDGS	EHSLEGGKFP	LEMQIYCFDA	DRPSSFEEAV	KKGKLRALS	180
60	ILFEVGTREN	LDKALIDGV	ESVSRFGKQA	ALDPFILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISESQL	AVPCEVLTMQ	QSGYVMLMDY	LQNNFREQQY	KFSRQVFSSY	300
	TGKEEIHAEV	CSSEPEENVQA	DPENYTSLLV	TWERPRVVDY	TMIEKPAVLY	QQLDGEDQTK	360
	HEFLTQGYQD	LQAILNLLNL	NMSYVLQIVA	ICTNGLYGYK	SDQLIVDMPT	DNPELDLFFE	420
	LIGTEEIIKE	EEBGKIDIEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTHYNYRI	GTKYNEAKTN	480
65	RSPTRGSEFS	GKGDVPTSL	NSTSQPVTKL	ATEKDISLTS	QVTLELPPT	VEGTSASLND	540
	QSKTVLRSPH	MNLSGTAESE	NTVSIETEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISQGYIFS	SENPEITTYD	VLIPESARNA	SEDSTSSGSE	ESLKDFSMEG	NWNFPSSTDI	660
	TAQPDVSGSR	ESFLQNTYTE	IRVDESEKTT	KSFAGPVMS	QGPSVTDLEM	PHYSTFAFYP	720
	TEVTPHAFTP	SSRQDLVST	VNVVYSQTTQ	PVYNBASNSS	HESRIGLAEG	LESEKKAVIP	780
70	LVIVSALTFI	CLVVLVGILI	YWEKCFQTAH	FYLEDSTSPR	VISTPPTPIF	PISDDVGAIP	840
	IKHFPKHVAD	LHASSGFTTE	FEEVQSCITVD	LGITADSSNH	PDNKHKNRYI	NIVAYDHSRV	900
	KLAQLAEKDG	KLTDYINANY	VDGYNRPKAY	LAAQGLKST	AEDFWRMIWE	HNVEVIVMIT	960
	NLVEKGRRC	DQYWPADGSE	EYGNFLVTQK	SVQVLAYYTV	RNFTLRNTKI	KKGSQKGRPS	1020
	GRVVYQYHYT	QWPDGMVPEY	SLFVLTFVRK	AAAYAKRHAVG	PVVVHCAGV	GRTGYIVLVD	1080
75	SMLQIQIHEG	TVNIFGLFKH	IRSQRNYLVQ	TEEQYVFIHD	TLVEALLSKE	TEVLDSHIHA	1140
	YVNALLIPGP	AGTKLEKQF	QLLSQSNIQQ	SDYSALKQOC	NREKNRTSSI	IPVERSRVGI	1200
	SSLSGEGTDY	INASYIMGY	QSNFIIITQH	PLLEHTIKDFW	RMIWDHNAQL	VVMIPDQNM	1260
	AEDEFVYWP	KDEPINCSEF	KVTLMAEEHK	CLSNEEKLII	QDFILEATQD	DYVLEVRHFQ	1320
	CPKWPNDSP	ISGTFELIS	IKEEAANRDG	PMIVHDEHGG	VTAGTFPCALT	TLMHQLEKEN	1380
80	SVDVYQVAKM	INLMRPGVFA	DIEQYQFLYK	VILSLVSTRQ	EENPSTSLDS	NGAALPDGNI	1440
	AESLESLV						

Seq ID NO: 185 DNA sequence  
Nucleic Acid Accession #: EOS sequence  
Coding sequence: 501-4514

85	1	11	21	31	41	51
----	---	----	----	----	----	----

	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAGCAG	AGGAGCCGCA	120
5	CGCGAGGGG	CCGCGAGCCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCT	CGCTTGCATT	180
	CAGCTCCTCT	GTGTTTCCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCTATACA	GGAGCACTGA	ATCAAAAAAT	TGGGGAAAGA	300
	AATATCCAAC	ATGTAATAGC	CCAAAACAA	CTCCTATCAA	TATTGATGAA	GATCTTACAC	360
	AAGTAAATGT	GAATCTTAAG	AAACTTAAT	TTCAGGGTTG	GGATAAAACA	TCATTGGAAA	420
10	ACACATTCAT	TCATAACACT	GGGAAAAAC	TGGAAATTA	TCTCACTAAT	GACTACCGTG	480
	TCAGCGGAGG	AGTTTCAGAA	ATGGTGTITA	AAGCAAGCAA	GATAACTTTT	CACTGGGGAA	540
	ANTCCATAT	GTCACTCGAT	GGATCAGAGC	ATAGTTTAGA	AGGACAAAA	TTTCCACTTG	600
	AGATGCCAAT	CTACTGCTTT	GATGCGGACC	GATTTTCAAG	TTTTGAGGAA	GCAGTCAAAG	660
	GAAGAGGGAA	GTTAAGAGCT	TTATCCATT	TGTTTGAGGT	TGGGACAGAA	GAATTTTGG	720
	ATTTCAAAGC	GATTATGTAT	GGAGTCGAAA	GTGTTAGTCG	TTTTGGGAAG	CAGGCTGCTT	780
15	TAGATCCATT	CATCTGTTTG	AACTTCTGCG	CAAACTCAAC	TGACAAGTAT	TACATTTACA	840
	ATGGCTCATT	GACATCTCCT	CCCTGCACAG	ACACAGTTGA	CTGGATTGTT	TTTAAAGATA	900
	CAGTTAGCAT	CTCTGAAGAC	CAGTTGGCTG	TTTTTTGTGA	AGTTCTTACA	ATGCAACAAT	960
	CTGGTTATGT	CATGCTGATG	GACTACTTAC	AAAACAAATT	TCCAGAGCAA	CAGTACAAGT	1020
20	TCTCTAGACA	GGTGTTTTCC	TCATACACTG	GAAGGAAGA	GATTCATGAA	CGAGTTTGT	1080
	GTTCAGAAC	AGAAAAATGT	CAGGCTGACC	CAGAGAATTA	TACCAGCCTT	CTGTGTACAT	1140
	GGGAAAGACC	TCCAGTCGTT	TATGATACCA	TGATTGAGAA	GTTTGAGT	TTGTACCAAGC	1200
	AGTTGGATGG	AGAGGACCAA	ACCAAGCATG	AATTTTGTAC	AGATGGCTAT	CAAGACTTGG	1260
	GTGCTATTCT	CAATAATTGG	CTACCCAATA	TGAGTTATGT	TCTTCAGATA	GTAGCCATAT	1320
25	GCACATAATG	CTTATATGGA	AAATACAGCG	ACCAACTGAT	TGTCGACATG	CCTACTGATA	1380
	ATCCTGAATC	TGATCTTTTC	CCTGAATTA	TGGAACTGA	AGAAATATC	AAGGAGGAGG	1440
	ANGAGGGAAA	GACCTTCTTA	GAAGGCGCTA	TTGTGAATCC	TGGTAGAGAC	AGTGCTACAA	1500
	ACCAATTCAG	GAAGAGGAAA	CCCCAGATT	CTACCAACA	ACACTACAAT	CGCATAGGGA	1560
	CGAAATACAA	TGAAGCCAAG	ACTAACCGAT	CCCCAACAG	AGGAAGTGAA	TTCTCTGGAA	1620
30	AGGGTGATGT	TCCCAATACA	TCTTTAAATT	CCACTTCCCA	ACCACTGACT	AAATTAGCCA	1680
	CAGAAAAAGA	TATTTCTTGG	ACTTCTCAGA	CTGTGACTGA	ACTGCCACCT	CACACTGTGG	1740
	AAGGTACTTC	AGGCTCTTTA	AATGATGGCT	CTAAACTGT	TCTTAGATCT	CCCATATGTA	1800
	ACTGTGCGGG	GACTGCGAGAA	TCCTTAAATA	CAGTTTCTAT	AACAGAAATAT	GAGGAGGAGA	1860
	GTTTATTGAC	CAGTTTCAAG	CTTGATCTG	GAGCTGAAGA	TTCTTCAGGC	TCCAGTCCCG	1920
35	CAACTCTGCG	TATCCCATTC	ATCTCTGAGA	ACATATCCCA	AGGGTATATA	TTTTCTCCCG	1980
	AAAACCCAGA	GACAAATAAC	TATGATGTCC	TTATACAGGA	ATCTGCTAGA	AATGCTTCCG	2040
	AGATTCAAC	TGATTCAGGT	TCAGAAAGAT	CACATAAGGA	TCCTTCTATG	GAGGGAAATG	2100
	TGTGGTTTCC	TAGCTCTACA	GACATAACAG	CACAGCCCGA	TGTTGGATCA	GGCAGAGAGA	2160
	GCTTTCTCCA	GACTAATTAC	ACTGAGATAC	GTGTTGATGA	ATCTGAGAG	ACAACCAAGT	2220
40	CTTTTCTGCG	AGGCCAGTG	ATGTCACAGG	GTCCCTCAGT	TACAGATCTG	GAATGCCAC	2280
	ATTATTCTAC	CTTTGCTTAC	TTCCCAACTG	AGGTAAACCC	TCATGCTTTT	ACCCATCCT	2340
	CCAGACAAAC	GGATTGGTTC	TCCACGGTCA	ACGTGGTATA	CTCGCAGACA	ACCCAACCGG	2400
	TATACATAG	GGCCAGTAAT	AGTAGCCATG	AGTCTCGTAT	TGGTCTAGCT	GAGGGGTTGG	2460
	AATCCGAGAA	GAAGGCAGTT	ATACCCCTTG	TGATCGTGTG	AGCCCTGACT	TTTATCTGTG	2520
45	TAGTGGTTCT	TGTGGGTATT	CTCATCTACT	GGAGGAAATG	CTTCCAGACT	GCACACTTTT	2580
	ACTTAGAGGA	CAGTACATCC	CCTAGAGTTA	TATCCACACC	TCCAACACCT	ATCTTTCCAA	2640
	TTTCAGATGA	TGTGGGAGCA	ATTCCAATTA	AGCACITTC	AAAGCATGTT	GCAGATTTAC	2700
	ATGCAAGTAG	TGGGTTTACT	GAAGAAATTT	AGACACTGAA	AGAGTTTAC	CAGGAAGTGC	2760
	AGAGCTGTAC	TGTTGACTTA	GSTATTACAG	CAGACAGCTC	CAACCAACCA	GACAAACAAG	2820
50	ACAAGAAATG	ATACATAAAT	ATCGTTGCC	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	2880
	TGCTGAAAA	GGATGGCAAA	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	2940
	ACAGACCAAA	AGCTTATATT	GCTGCCCAAG	GCCCACTGAA	ATCCACAGCT	GAAGATTCT	3000
	GGAGAAATAT	ATGGGAAACAT	AATGTGGAAG	TTATTGTCTAT	GATAACAAAC	CTGTTGGAGA	3060
	AAGGAAGGAG	AAAGTGTGAT	CAGTACTGGC	CTGCCGATGG	GAGTGAGGAG	TACGGGAAC	3120
55	TTCTGGTTC	TCAGAGAGGT	GTGCAAGTGC	TTGCTTATTA	TACTGTGAGG	AATTTTACTC	3180
	TAAGAAACAC	AAAAATAAAA	AAGGCTCC	AGAAAGGAAG	ACCCAGTGG	CGTGTGGTCA	3240
	CACAGTATCA	CTACAGCCAG	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	CTGCTGCTCC	3300
	TGACCTTTGT	GAGAAAGGCA	GCCTATGCCA	AGGCCCATGC	AGTGGGGCT	GTGCTGCTCC	3360
	ACTGCAATGC	TGTTGTTGGA	AGAAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	3420
60	AGATTCAACA	CGAAGGAAT	GTCAACATAT	TTGGCTTCTT	AAAAACATC	CGTTCAACAA	3480
	GAATATTATT	GGTACAACT	GAGGAGCAAT	ATGCTTCTAT	TCATGATACA	CTGGTTGAGG	3540
	CCATACTTAG	TAAAGAACT	GAGGTGCTGG	ACAGTCAAT	TCATGCTAT	GTTAATGCAC	3600
	TCTCATTC	TGGACAGCA	GGCAAAACAA	AGCTAGAGAA	ACAATTCAG	CTCCTGAGCC	3660
	AGTCAATAT	ACAGCAGAGT	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAAAGA	3720
65	ATCGAACTTC	TTCTATCATC	CCTGTGGAAA	GATCAAGGGT	TGGCAITTC	TCCCTGAGTG	3780
	GAGAAAGCAC	AGACTACATC	AATGCCTCCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	3840
	TCATCATTAC	CCAGCACCTC	CTCCTTCATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	3900
	ACCATATGCG	CCAATGGTG	GTTATGATTC	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	3960
	TTGTTTACTG	GCCAAATAAA	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACTCTTA	4020
70	TGGCTGAAGA	ACACAAATGT	CTATCTAATG	AGGAAAAACT	TATAATTTCAG	GACTTTATCT	4080
	TAGAAGCTAC	ACAGGATGAT	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CCTAAATGGC	4140
	CAATCCAGA	TAGCCCATTT	AGTAAACTT	TTGAACITAT	AAGTGTATTA	AAAGAAGAAG	4200
	CTGCCAATAG	GGATGGGCTT	ATGATTGTTC	ATGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	4260
	CTTTCTGTGC	TCTGACAAAC	CTTATGCACC	AACTAGAAAA	AGAAAAATTC	GTGGATGTTT	4320
75	ACCAGGTAGC	CAAGATGATC	AATCTGATGA	GGCCAGGAGT	CTTTGCTGAC	ATTGAGCAGT	4380
	ATCAGTTTCT	CTACAAAGTG	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAAATCCAT	4440
	CCACCTCTCT	GGACAGTAAT	GGTGAGCAT	TGCCGTATGG	AAATATAGCT	GAGAGCTTAG	4500
	AGTCTTTAGT	TTAACACAGA	AAGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTCTCTCT	4560
	TCCTAAAAAT	AGGAGGAAA	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTT	CCATCACCTG	4620
80	ACAGTAACCT	TCATGACATA	GGATTCTGCC	GCCAAATTTA	TATCATTAAC	AATGTGTGCC	4680
	TTTTTGCAG	ACTTGTAAAT	TACTATTAT	GTTTGAACCT	AAATGATTGA	ATTTTACAGT	4740
	ATTTCTAAGA	ATGGAATTTG	GGTATTTTT	TCTGTATTGA	TTTTAACAGA	AAATTTCAAT	4800
	TTATAGAGGT	TAGGAATTC	AAACTACAGA	AAATGTTTGT	TTTTAGTGT	AAATTTTATG	4860
	CTGTATTGT	AGCAATTATC	AGGTTTGCTA	GAATATAAC	TTTTAATACA	GTAGCCTGTA	4920
85	AAATAAACAC	TCTTCCATAT	GATATTCAAC	ATTTTACAAC	TGCAGTATTC	ACCTAAAGTA	4980
	GAATAAATCT	TGTACTTAT	GTAAATACTG	CCCTAGTGTG	TCCATGGACC	AAATTTATAT	5040
	TTATATTTGT	AGATTTTAT	ATTTTACTAC	TGAGTCAAGT	TTTCTAGTTC	TGTGTAAATG	5100
	TTTAGTTTAA	TGACGTAGTT	CATTAGCTGG	TCTTACTCTA	CCAGTTTCT	GACATTTGAT	5160

WO 02/086443

PCT/US02/12476

TGTGTTACCT AAGTCATTAA CTTTGTTCAT GCATGTAATT TTAACCTTTTG TGGAAATAG 5220  
 AAAATACCTTC ATTTTGAAG AAGTTTTTAT GAGAATAACA CCTTACCAAA CATTGTTCAA 5280  
 ATGGTTTTTA TCCAAGGAAT TGCAAAAATA AATATAATA TTGCCATTAA AAAAAAATA 5340  
 AAAAAAATA AAAAAAATA AAA

Seq ID NO: 186 Protein sequence:  
 Protein Accession #: EOS sequence

1	11	21	31	41	51	
MVFKASKITF	HWGKCNMSSD	GSEHSLEGQK	FPLEMQIYCF	DADRFSSPFE	AVKKGKGLRA	60
LSILFEVGTB	ENLDFKAIID	GVESVSRFGK	QAALDPFILL	NLLPNSTDKY	YIYNGSLTSP	120
PCDTVDVNI	PKDTVSISES	QLAVFCEVLT	MQQSGYVMLM	DYLQNNFREQ	QYKFSRQVFS	180
SYTGKEEIH	AVCSSEPENV	QADPENYISL	LVTWERPRVV	YDTMIEKFAV	LYQLDGEDQ	240
TKHEFLTDGY	QDLGAILANL	LPNMSYVLQI	VAICTNGLYG	KYSDQLIVDM	PTDNEPELDF	300
PELIGTEBII	KEEKEGKDIE	EGAIIVNPRD	SATNQIRKKE	PQISTTTHYN	RIGTKYNEAK	360
TRNSPTRGSE	PSGKADVPNT	SLNSTSQPVT	KLATEKDIDL	TSQTVTELPP	HTVEGTSASL	420
NDGSKTVLR	PHMNLSGTAE	SLNTVSIET	EEESLLTSFK	LDTGAEDSSG	SSPATSAIPP	480
ISENISQGYI	FSSENPETIT	YDVLIPESAR	NASEDSTSSG	SEESLKDPSP	EGNWFPPSST	540
DITAQPDVGS	GRESFLQTNV	TEIRVDESEK	TKSFSAGFV	MSQGPSTVDL	EMPHYSTFAY	600
FPTEVTPEAF	TPSSRQQLDV	STVNVVYSQT	TQPVVNEASN	SSHESTRIGLA	EGLESEKKAV	660
IPLVIVSALT	FICLVVLVGI	LIYWRKCFQT	AHFYLEDSTS	PRVISTPTPT	IFPISDDVGA	720
IPKHFPPKHV	ADLHASSGFT	EEFETLKEFY	QEVQSCITDL	GITADSSNHP	DNKHKNRYIN	780
IVAYDHSRVT	LQALEKDGK	LTDYINANYV	DGYNRPKAYI	AAQGPLKSTA	EDFWRMWIEH	840
NVEVIMVTN	LVEKGRRKCD	QYWPADGSEB	YGNFLVTQKS	VQVLAYTVTR	NFTLNTKIK	900
KGSQKGRPSG	RUVTYQHYTY	WPDGMVPEYS	LPVLTFRVKA	AYAKRHAVGP	VVHCSAGVG	960
RTGTIVILDS	MLQQLQHEGT	VNIFGFLKHI	RSQRNYLVQT	EBQYVFIHDT	LVEALLSKET	1020
EVLDSHIBAY	VNALLIPGPA	GKTKLEKQFP	LLSQSNIQQS	DYSALKQCN	REKNRTSII	1080
PVERSRLGIS	SLSGEGTYI	NASYIMGYQ	SNEFIITQHP	LLETIKDFWR	MIWDHNAQLV	1140
VIMPDPGMA	DEPFYVWPK	DEPINCESFK	VTLMABEHC	LSNEEKLIQ	DFILEATQDD	1200
YVLEVRHPQC	PKWPNPDSPI	SKTFELISVI	KEEAANRDP	MIVHDEHGGV	TAGTFCALT	1260
LMHQLEKENS	VDVYQVAMI	NLMRPGVFAD	IBQYQFLYKV	ILSLVSTRQE	ENPSTSLDSN	1320
GAALPDGNIA	ESLESILV					

Seq ID NO: 187 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4632

1	11	21	31	41	51	
CACACATACG	CACGACGAT	CTCACTTGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
CAAAAAAAC	ATTTCTCTCG	CTCCCTCTCC	CTCTCCACTC	TGAGAGCAG	AGGAGCCGCA	120
CGCGAGGAGG	CGCGACACCG	TCTGGAAATG	CGAATCTCTA	AACGTTTCCT	CGCTTGCAAT	180
CAGCTCTCTCT	GTGTTTCCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
CTTGTGGAAG	AGAGTTGGCTG	GTCCTATACA	GAGAGCACTGA	ATCAAAAAAA	TGGGGGAAAG	300
AAATATCCAA	CATGTAAATG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTTGAA	420
AACACATCTA	TTCTATAACAC	TGGGAARACA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
GTCAAGGAGG	GAGTTTCAGA	AATGTTGTTT	AAAGCAAGCA	AGATAACTTT	TCATCGGGGA	540
AAATGCAATA	TGTCACTCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
GAGATGCAAA	TCTACTGCTT	TGATCGGAC	CGATTTTCAA	GTTTGTAGGA	AGCAGTCAAA	660
GGAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTGGGAA	GCAGGCTGCT	780
TTAGATCCAT	TCATACTGTT	GAACTCTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTAC	840
AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTGTGTTG	AAGTTCTTAC	AATGCAACAA	960
TCGTGTTATG	TCATGCTGAT	GGAATCTTAA	CAAAACAAAT	TTGAGAGCA	ACAGTACAAG	1020
TTCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
AGTTCTAGAC	CAGAAATAGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCCTGTTTACA	1140
TGGGAAAGAG	CTGAGATCGT	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCA	1200
CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
GGTGCTATT	CAATAAATTT	GCTACCCAA	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
AATCTCTGAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACCT	AAGAAATAAT	CAAGGAGGAG	1440
GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
AACCAATATCA	GGAAAGAGGA	ACCCAGAGTT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACAGTCAC	TAAATTAGCC	1680
ACAGAAAGAG	ATATTTCCCT	GACTTCTCAG	ACTGTGACTG	AACAGCCACC	TCACACTGTG	1740
GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCACATATG	1800
AACCTGTGCG	GGAGCTGCAG	ATCCTTAAAT	ACAGTTTCTA	TAAAGAAATA	TGAGGAGGAG	1860
AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
GCAACTCTCT	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGTATAT	ATTTCTCTCC	1980
GAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCA	AATCTGCTAG	AAATGCTTCC	2040
GAAGATTCAA	CTTCACTCAG	TTCAGAGGAA	TCCTTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
GTGTGTTTCT	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTTGATC	AGGCAGAGAG	2160
AGCTTTCTCC	AGACTAATTA	CAGTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
TCCTTTCTCT	CAGGCCAGT	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
CATTATTCTA	CTTTGCGCTA	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
TCCAGACAAC	AGGATTGTTG	CTCCACGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCCAACG	2400
GTATACAAAT	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACTCTTT	2580
TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
ATTTTCAGATG	ATGTCGGAGC	AATTCCAATA	AAGCACTTTC	CAAAGCATGT	TGCAGATTTA	2700
CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	2760
CAGAGCTGTA	CTGTGACTTT	AGGTATTACA	GCAGACAGCT	CCAACCAACC	AGACAACAAG	2820

5  
10  
15  
20  
25  
30  
35  
40  
45

CACAAGAATC GATACATAAA TATCGTTGCC TATGATCAT GACAGGTAA GCTAGCACAG 2880  
 CTGCTGAAA AGGATGGCAA ACTGACTGAT TATATCAATG CCAATTATGT TGATGGCTAC 2940  
 AACAGACCAA AAGCTTATAT TGCTGCCCAA GGCCCACTGA AATCCACAGC TGAAGATTTC 3000  
 TGGAGAAATGA TATGGGAACA TAATGTGGAA GTTATTGTCA TGATAACAAA CCTCGTGAG 3060  
 AAAGGAAGGA GAAATGTGA TCAGTACTGG CCTGCCGATG GGAGTGAGGA GTACGGGAAC 3120  
 TTTCTGTGCA CTCAGAAGAG TGTGCAAGTG CTGCTATT ATACTGTGAG GAATTTTACT 3180  
 CTAAGAAACA CAAAATAAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACCTGTGGTC 3240  
 ACACAGTATC ACTACACGCA GTGGCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 3300  
 CTGACCTTTC TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTGTGCTC 3360  
 CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGACG 3420  
 CAGATTCAAC AACTGAAGAAC TGTCACATA TTTGGCTTCT TAAACACAT CGTTTCAAA 3480  
 AGAAATTATT TGGTACAAAC TGAGGAGCAA TATGCTTCA TTCATGATAC ACTGGTTGAG 3540  
 GCCATCTTAT GTAAAGAAAC TGAGGTGCTG GACAGTATA TTCATGCCCTA TGTTAATGCA 3600  
 CTCTCATTC CTGGACCAAG AGGCAAAACA AAGCTAGAGA AACAATTCCA GGGTCTCACT 3660  
 CTGTACCCCA GGTGGAGTGG CAGAGGCACA ATCTCGGCTC ACTGCAACCT TCTCTCCCT 3720  
 GGCCTTAAGT ATCTCTCTAC CTGAGCTCC CGAGTGGCTG GGACTATACT CCTGAGCCAG 3780  
 TCAATATATC AGCAGAGTGA CTATTCTGCA GCCTAAAGC AATGCAACAG GGAAGAAGAT 3840  
 CGAATCTTCT CTATCATCCC TGTGGAAGA TCAAGGTTG GCATTTTATC CCTGAGTGGG 3900  
 GAAGGCACAG ACTACATCAA TGCTCTCTAT ATCATGGGCT ATTACAGAG CAATGAATTTC 3960  
 ATCATTACCC AGCACCCTCT CCTTCATACC ATCAAGGATT TCTGAGGAT GATATGGGAC 4020  
 CCAATATGCC AACTGGTGGT TATGATTCCT GATGGCCAAA ACATGGCAGA AGATGAATTTC 4080  
 GTTTACTGGC CAAATAAGA TGAGCCTATA AATTGTGAGA GCTTTAAGST CACTCTTATG 4140  
 GCTGAAGAAC ACAAATGCTC ATCTAATGAG GAAAACTTA TAATTGAGG CTTTATCTTA 4200  
 GAAGCTACAC AGGATGATTA TGTACTTGAA GTGAGGCACT TTCAGTGTCC TAAATGGCCA 4260  
 AATCCAGATA GCCCATTAG TAAACCTTTT GAACCTATAA GTGTATATAA AGAAGAAGCT 4320  
 GCCAATAGGG ATGGGCTCAT GATTGTTTAT GATGAGCATG GAGGAGTGAC GGCAGGAAGT 4380  
 TTCTGTGCTC TGACCAACCT TATGCACCAA CTAGAAAAG AAAATTCGAT GGTGTTTAC 4440  
 CAGGTAGCCA AGATGATCAA TCTGATGAG CCAGGAGTCT TTGCTGACAT TGAGCAGTAT 4500  
 CAGTTTCTCT ACAGAGTGAT CCTCAGCCTT GTGGGCACAA GGCAGGAAGA GAATCCATCC 4560  
 ACCCTCTCTG ACAGTATGAG TGCAGCATG CCTGATGGAA ATATAGCTGA GAGCTTAGAG 4620  
 TCTTTAGTTT AACACAGAAA GGGGTGGGG GACTCACATC TGAGCATTTG TTTCTCTTC 4680  
 CTAATATTAG GCAGGAAAT CAGTCTAGTT CTGTTATCTG TTGATTTCCC ATCACCAGAC 4740  
 AGTAACCTTC ATGACATAGG ATCTGCCGCG CAATTTTATA TCATTACAA TGTGTGCTT 4800  
 TTTGCAAGAC TTGTAATTTA CTATTATGT TTGAACATAA ATGATTGAAT TTTACAGTAT 4860  
 TTCTAAGAT GGAATGTGG TATTTTTC TGTATTGATT TTAACAGAAA ATTTCAATT 4920  
 ATAGAGGTTA GGAATTCCTA ACTACAGAAA ATGTTTGT TTAGTGTCAA ATTTTAGCT 4980  
 GTATTGTAG CAATTATCAG GTTTGTGAGA AATATACTT TTAATACAGT AGCCTGTAAA 5040  
 TAAACACTC TTCCATATGA TATTCACAT TTTACACTG CAGTATTCAC CTAAGTAGA 5100  
 AATAATCTGT TACTTATGT AAATACTGCC CTAGTGTCTC CATGACCCTA ATTTATATT 5160  
 ATAATGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTTCTG TGTAAATGTT 5220  
 TAGTTTAAAG ACGTAGTTCA TTAGCTGGTC TTACTCTACC AGTTTCTGCA CATTGTATTG 5280  
 TGTTACCTAA GTCAATTAAT TTGTTTCTG ATGTAATTTT AACTTTTGTG GAAATAGAAA 5340  
 ATACCTTCAT TTTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TTTTCAAT 5400  
 GGTTTTTATC CAAGGAATTG CAAAAATAA TATAAATATT GCCATTAATA AAAAAAATA 5460  
 AAAAAAATA AAAAAAATA A

Seq ID NO: 188 Protein sequence:  
 Protein Accession #: EOS sequence

50  
55  
60  
65  
70  
75

1 11 21 31 41 51  
 MRILKRFLAC IQLLCVCRLD WANGYYRQOR KLVEEIGWSY TGALEKQKNG KRYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKRL KFGQWDXLSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFWH GKCNMSSDGS EHSLEGGKFP LEMQIYCFDA DRFSFEEAV KGKGLRLALS 180  
 ILFEVGTEN LDFKAIIDGQ ESVSFRGKQA ALDPFILLNL LPNSTDKYIY YNGSLTSPPC 240  
 TDTVDVIVFK DTVSISBSQL AVFCEVLTMO QSGYVLMMDY LQNNFREQQY KFSRQVFSY 300  
 TGKEEIEHAV CSSEFENVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGDQTK 360  
 HEPLTDGQD LGALLNLLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNPELDLPE 420  
 LIGTESIIEE EEEGKDIEEG AIVNPRGDSA TNQIRKKEPQ ISTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGQVNTSL NSTSQPVTKL ATEKDISLTS QTVTELPHT VEGTSASLND 540  
 GSKTVLRSPH MNLSTABSL NTVSITYEES ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGVIPS SENPETITYD VLPESARNA SEDSTSSGSE ESLKDPMSMG NVWFPSTDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSPSAGPVMS QGSPVTDLEM PHYSTFAYFP 720  
 TEVTPHAFTP SSRQODLVST VNVVYSQTTO FVYNEASNSS HESRIGLAGS LESEKKAVIP 780  
 LVIVSALTPI CLVVLVGLI YWRKCPQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840  
 IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCVTLGI TADSSNHPDN KHKRYINIV 900  
 AYDHSRVKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA QGPLKSTAE PWRMIWEHV 960  
 EVIVMITNLV EKGRRKCDQY WPADGSEYQ NFLVTQKSVQ VLAYTTRNF TLRNTKIKKG 1020  
 SQKGRPSGRV VTOYEYQWPM DMGVPEYSLP VLTFRKAAY AKRHAVGPV VHCAGVGRT 1080  
 GTYIVLDSML QQIQHEGTVN IFGLKHRS QRNVLVQTEE QYVFIHDTLV EAILSKETE 1140  
 LDSHIIHAYN ALLIPGPAGV TKLEKQFQGL TSPRLCERG TISAHCNLPL PGLTDPPTSA 1200  
 SRVAGTILLS QSNIQSDYS AALKQCNREK NRTSSIIPE RSRVGISLS GEGTDYINAS 1260  
 YIMGYYSNE FIITQHPLP TIKDFWRMIW DINAQLVVM PDGQNMAGE FVYWPKNDEP 1320  
 INCSPKVTL MAEEHKCLSN BEKLIQDFI LEATQDDYVL EVRHFCQPKW PNPDSPIST 1380  
 FELISVKEE AANRDGPMIV HDEHGGVTAG TFCALTLMH QLEKENSVDV YQVAKMINLA 1440  
 RFGVFADIEQ YQFLYKVILS LVGTQREENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 189 DNA sequence  
 Nucleic Acid Accession #: NM\_002820  
 Coding sequence: 304..831

85

1 11 21 31 41 51  
 CCGGTTGCGA AAGAAGCTGA CTTGAGAGGG GAAACCTTC TTTCTTTAGG AGGCGGTTAG 60  
 CCTGTGTTCA CGAACCCAGG AGAAGTCTG GCCAGATTAA TTAGACATTG CTATGGGAGA 120  
 CGGTAAACA CACTACTTAT CATTGATGCA TATATAAAC CATTTTATT TCGCTATTAT 180

TTCAGAGGAA GCGCCTCTGA TTGTTCCTTT TTTCCTCTTC TGGCTGTGTG 240  
 GTTTGGAGAA AGCAGACTGT GAGTAGCCGG TTGCTAAATA AGTCCCGAGC GCGAGCGGAG 300  
 ACGATGCAGC GGAGACTGGT TCAGCAGTGG AGCGTCCGCG TGTTCTCTGCT GAGCTACGCG 360  
 GTGCCCTCCT GCGGGCGCTC GGTGGAGGGT CTCAGCCGCC GCCTCAAAAG AGCTGTGTCT 420  
 GAACATCAGC TCCTCCATGA CAAGGGGAAG TCCATCCAG ATTACGGCG ACGATTCTTC 480  
 CTTCACTATC TGATGCGAGA AATCCACACA GCTGAAATCA GAGCTACCTC GGAGGTGTCC 540  
 CCTAATCCA AGCCTCTCTC CAACACAAAG AACCACCCCG TCCGATTGGG GTCTGATGAT 600  
 GAGGGCAGAT ACCTAACTCA GGAACCTAAC AAGGTGGAGA CGTACAAAGA GCAGCCGCTC 660  
 AAGACACCTG GGAAGAAAAA GAAAGGCAAG CCGGGGAAC GCAAGGAGCA GGAAAGAAA 720  
 AAACGCGGAA CTGCTCTGCT CTGGTTAGAC TCTGGAGTGA CTGGGAGTGG GCTAGAAGGG 780  
 GACCACCTGT CTGACACCTC CACAACGTCG CTGGAGCTCG ATTACGGTA ACAGGCTTCT 840  
 CTGGCCCGTA GCGTCAGCGG GGTGCTCTCA GCTGGGTTTT GGAGCCTCCC TTCTGCCTTG 900  
 GCTTGGACAA ACCTAGAATT TTCTCCCTTT ATGTATCTCT ATGATTGTG TAGCAATTGA 960  
 CAGAGAATAA CTAGAATAT TGTCTGCCTT AAAGCAGTAC CCCCCTACCA CACACACCCC 1020  
 TGTCTCCAG CACCTAGAG AGGCGCTAGA GCCCATTCCT CTTTCTCCAC CGTACCCCAA 1080  
 CATCAATCTT TTACCACTCT ACCAATAAT TTCAATTCA AGCTTCAGAA GCTAGTGACC 1140  
 ATCTTATAA TTGTCTGAG AGGTGTATTT CTCCCTCTA CTCTCACACC TGGGCAAACT 1200  
 TTCTTCAAG TTTTTCATT CTACCTTCT TCACTTCAA GGGAGAATAT AGAAGCATTT 1260  
 GATATTATCT ACAACACTG CAGAACAGCA TCATGTCTA AAGATTCTG AGCCATTAC 1320  
 ACTTTTATTT TAATAAATG TATTTAATTA AATCTCAAT TTATTTTAA GTAAAGAACT 1380  
 TAAATTATCT TTTAAACCA TGCCTTAAAT TTGTTTAAT AAATTTAAT CTGTTTCTA 1440  
 CCACTCATA CAAATAAAT GGTTCCTGAA AATGTTAAG TATTAACCTA CAGGATATA 1500  
 GGTTTTCTC ATGTATCTTT TTGTTCAATG GCAAGATGAA ATAAATTTTC TAGGGTAATG 1560  
 CCGTAGGAAA AATAAACTT CACATTAAA AAAAA

Seq ID NO: 190 Protein sequence:  
 Protein Accession #: NP\_002811

1 11 21 31 41 51  
 MQRRLVQWVS VAVFLLSYAV PSCGRSVEGL SRRLKRAVSE HQLLHDKGKS IQDLRRRFPL 60  
 HELIAEHTA EIRATSEVSP NSKPSFNTKN HPVRFGSDDE GRVLTQETNK VETYKEQPLK 120  
 TPGKKKGKPK GKRRKQEKKK RRTRSAWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 191 DNA sequence  
 Nucleic Acid Accession #: XM\_059328  
 Coding sequence: 52..1023

1 11 21 31 41 51  
 GGGCTGTCCG GCGCACTCCC CTGGGAGCGC GAGCGGTGGA CCCAGGCGGC CATGTCCCGC 60  
 CCTCGCATGC GCGTGTGGT CACCGCGGAC GACTTTGGTT ACTGCCCGCG ACGCGATGAG 120  
 GGTATCGTGG AGGCTTTTCT GCGCGGGGCT GTGACCAGCG TGTCCTCTGCT GGTCAACGGT 180  
 GCGGCGACCG AGAGCGCGGC GGAGCTGGCC CGCAGGCACA GCATCCCCAC GGGCTCCAC 240  
 GCCAAGCTGT CCGAGGGCGC CCGCGTGGGT CCGGCGCGCT GTGGCGCTC ATCGCTGCTC 300  
 GCGCGGAAAG GCTTCTTCTT TGGCAAGATG GGATTCCGGG AGGCGGTGGC GCGCGGAGAC 360  
 GTGGATTGCG CTACGGTGGC GGAGGAGCTC GAGGCCCAAC TAAGCTGCTT CCGGAGCTG 420  
 CTGGGCGGG CCCCCAGCA CGCGGACGGG CACGAGCAGG TGCACTGCTT CCCAGCGGTG 480  
 TGCCAGGTGT TCGCGAGGC GCTGAGGCC TATGGGTGTC GCTTTACGCG ACTGCCGCTG 540  
 GAGCGCGGTG TGGGTGGCTG CACTTGGCTG GAGGCCCGCG CGGTGCGCTT CGCTGCGCC 600  
 GTGGAGCGCG AGCGCGGGC CGCGTGGGC CCCTTCTCCC GCCACGGCTT GCGGTGGACA 660  
 GACGCTCTCG TGGGCTGAG CACTTGGCGC CGGCACATGT CGCTCACCG CGTGTCCGGG 720  
 GCGCTGGCG GGTCTCTGGA AGGTACCTTA GCGGCCACA CCCTGACAGC CGAGCTGATG 780  
 GCGCACCCCG GCTACCCAG TGTGCTCCC ACGCGCGGCT GCGGTGAAGG CCCCAGCGCT 840  
 TTCTCTGCT CTGGGAGCG GCTGCATGAG CTGCGCTGCC TCACCGCGCT CACGCTGCGG 900  
 GCCAGCTTG CCGAGGATGG CGTGAGCTT TGGCGCTCG ACGACCTGGA CTCGAGAGG 960  
 CCAAGGGAGG AGGTCCCCTG TGAGCCCACT CTGGAACCTT TCCTGGAACC CTCCTACTC 1020  
 TGACCCCTTA CAGACAACA AGCACTAATC CCCTTAGTAC CAAGAAAGGG GAGCCAGGAT 1080  
 TTAGTCTTGG CCCAGCCAG AGCTGGGACC TGGAGCAGA TCTGTGACT TCCCTGGGTA 1140  
 GGACACTGCC ACCTCTGGG TCAGGTCTCT ATGCTTCAA ATGGCATCTA GAGTTTGAGC 1200  
 AGCCTTCTTG GCTGCAGGCA GGCCTAGCCT GTGGCAGCG GCTAGGCGCC GCAGAGCATT 1260  
 TGGTGGCCTT CCATGTTGCA ATGCAACAC CTTCAACCT GGGGAGTGG GGAGAGATGG 1320  
 CTATATTAAT AAAATAACGT GTGTCTTTC

Seq ID NO: 192 Protein sequence:  
 Protein Accession #: XP\_059328

1 11 21 31 41 51  
 MSRPRMLVV TADDFGYCPR RDEGIVEAFL AGAVTSVSLV VNGAATESAA ELARRHSIPT 60  
 GLHANLSEGR PVGPARRGAS SLLGPEGFFL GRMGFREAVA AGDVLDPQVR EELEAQLSCF 120  
 RELLGRAPTH ADGHQHVHVL PGVCQVFAEA LQAYGVRFTR LPLERGVGCC TWLEAPARAF 180  
 ACAVERDARA AVGPFSRHGL RWTDAFVGLS TCGRHMSAHR VSGALARVLE GTLAGHTLTA 240  
 ELMAHPGYFS VPPTGGCGEG PDAFSCSWER LHELRLVTAP TLRAQLAQDG VQLCALDDLD 300  
 SKRPGEEVPC EPTLEPFLEP SLL

Seq ID NO: 193 DNA sequence  
 Nucleic Acid Accession #: NM\_005688.1  
 Coding sequence: 126..4439

1 11 21 31 41 51  
 CCGGGCAGGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCTGGAGC 60  
 AGGGGCGCAG GAATCTGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCGCTCAG 120



	AGAAGATGAA	GGATATCGAC	ATAGGAAAAG	AGTATATCAT	CCCCAGTCTT	GGGTATAGAA	180
	GTGTGAGGGA	GAGAACCCAG	ACTTCTGGGA	CGCACAGAGA	CCGTGAAGAT	TCCAGTTCA	240
	GGAGAACTCG	ACCGTTGGAA	TGCCAAGATG	CCTTGGAAAC	AGCAGCCCGA	GCCAGGGGCC	300
5	TCTCTCTTGA	TGCTCTCATG	CATTCTCAGC	TCAGAATCCT	GGATGAGGAG	CATCCCAAGG	360
	GAAAGTACCA	TCATGGCTTG	AGTGTCTCTA	AGCCCCATCG	GACTACTTCC	AAACACCAGC	420
	ACCCAGTGGG	CAATGCTGGG	CTTTTTCTCT	GTATGACTTT	TTCTGGCTTT	TCTTCTCTGG	480
	CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
10	AAGTTGGGCC	AGACGCTGCT	TCCCTGCGAA	GGGTGTGTGT	GATCTTCTGC	CGCACCAGGC	660
	TCATCTCTGC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAGCCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTTAGT	GCTGGGCTTC	CTCCTGACGG	AAATCGTGGC	GTCTTGGTGG	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCGA	ACCGGTGTCC	GCTTGGGGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	TACTTAAGTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
15	TTTGCTCCAA	CGATGGGCGA	AGAAATGTTG	AGGCAGCAGC	CGTTGGCAGC	CTGCTGGCTG	1020
	GAGGACCCGT	TGTTGCCATG	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCTCTGG	ATCAGCTGTT	TTTATCTCTT	TTTACCCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CCGCCACCGA	TGAACGTGTC	CAGAAGATGA	1200
20	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTTCTC	1260
	AGAGTGTTC	AAAAATCCGC	GAGGAGGAGC	GTCCGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGATACAC	TGTTGGGTGT	GCTCCCATTG	TGGTGGTGTG	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTCAATAT	GACCTCTGGC	TTCCGATCTG	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	TCTTCAATTC	CATGACTTTT	GCTTTGAAAG	TAACACCGTT	TTCACTAAAG	TCCCTCTCAG	1500
25	AAGCCTCAGT	GGCTGTTGAC	AGATTAAAGA	GTITGTTTCT	AATGGAAGAG	GTTCACATGA	1560
	TAAAGAACAA	ACCAAGCCAGT	CCTCACATCA	AGATAGAGAT	GAAAAATGCC	ACCTTGGCAT	1620
	GGGACTCTTC	CCACTCCAGT	ATCCAGAACT	CGCCCCAGCT	GACCCCCAAA	ATGAAAAAAG	1680
	ACAAGAGGGC	TTCCAGGGGC	AAGAAGAGAG	AGGTGAGGCA	GCTGAGCGCG	ACTGAGCATC	1740
	AGGCGGTGCT	GCCAGAGCAG	AAAGGCCACC	TCCTCCTGGA	CAGTGAAGAG	CGGCCCACTG	1800
30	CCGAAGAGGA	AGAAGGCAAG	CACATCCACC	TGGGCCACCT	GCGCTTACAG	AGGACACTGC	1860
	ACAGCATCGA	TCTGGAGATC	CAAGAGGGTA	AATCTGGTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCAAT	TCAGCCATTT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTTC	GCTTATGTGG	CCCAGCAGGC	CTGGATCCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
35	ACAGCTGCTG	CCTGAGGCCT	GACCTGGCCA	TTCTTCCCAG	CAGCGACCTG	ACGGAGATTG	2160
	GAGAGCGAGG	AGCCCAACCTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGGCCT	2220
	TGTATAGTGA	CAGAGAGCAT	TACATCCTGG	ACGACCCCTC	CAGTGCTTGA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTTACCCA	CAAGTTACAG	TACCTGGTGG	ACTGTGATGA	AGTGAATCTT	ATGAAAGAGG	2400
40	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AATGATGAA	TTTAAATGGT	GACTATGCTA	2460
	CCATTTTATA	TAACTGTGTG	CTGGAGAGAG	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCAAGTG	TCTACAGAA	AAATCAACA	ACAAGGCTCC	TAAACACAGG	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAAGTCT	CTGGTCACTA	TATGGTGTCT	ACATCCAGGC	TGCTGGGGCG	CCCTTGGCAT	2700
45	TCTGTGTTAT	TATGGCCCTT	TTCTATGCTA	ATGTAGGCAG	CACCGCCTTC	AGCACCTGGT	2760
	GGTTGAGTTA	CTGGATCAAG	CAAGGAAGCG	GGAAACCCAC	TGTGACTCGA	GGGAACGAGG	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACAATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCTCTA	AAGCCATTCG	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCT	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTT	CCGAAGGATC	CTTCGAAGCC	3000
50	CTATGAAGTT	TTTTGACACG	ACCCCCACAG	GGAGGATTCT	CAACAGGTTT	TCCAAAGACA	3060
	TGGATGAAGT	TGAAGTGGCG	CTGCCGTTC	AGGCGAGAT	GTTCATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCCT	TTGGCAGTGG	3180
	GGCCCTCTGT	CATCCTCTTT	TCAGTCTCTG	ACATTGTCTC	CAGGCTCCTG	ATTCTGGAGC	3240
	TGAAGGTCT	GGACAATATC	ACGCAATCAC	CTTTCTCTCT	CCACATCAGC	TCCAGCATAC	3300
55	AGSGCCTTGC	CACCATCCAC	GCCTACAATA	AAGGGCAGGA	GTTCCTGCAC	AGATACCAGG	3360
	AGCTGTCTGA	TGACAACCAA	GCTCCTTTTT	TTTTGTTTAC	GTGTGCGATG	CGTGTGGCTG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCC	CACGGGGCTG	ATGATGGTTC	3480
	TTATGACAGG	GCAGATTCCC	CCAGCCTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
	TAAAGGGGCT	GTCCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTCACTT	3600
60	CGGTGGAGAG	GATCAATCAC	TACATTAAAG	CTCTGTCTCT	GGAGGACACT	GCCAGAATTA	3660
	AGAACAAAGC	TCCCTCCCTC	GACTGGCCCC	AGGAGGAGAG	GGTGACCTTT	GAGAACGCGG	3720
	AGATGAGGTA	CCGAGAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAA	3780
	CTAAGAGAGA	GATTGGCAAT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTCTCG	TCTGGTGGAG	TTATCTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
65	GTGATATTGG	CCTTGGCGAC	CTCGAAGCA	AATCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
	TGTTCACTGG	CATGTCAGAG	TCAAAATTGG	ACCCCTTCAA	CCAGTACACT	GAAGACCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CCTCTGAAAC	4080
	TGAATCTGA	AGTGATGGAG	AATGGGATA	ACTTCTCAGT	GGGGGAACGG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTATGATGA	GCCACAGCTG	4200
70	CCATGACAC	AGAGACAGAC	TTATTGATTC	AAGAGACCAT	CCGAGAGACA	TTTGACAGCT	4260
	GTACCATGCT	GACCATTGCC	CATCGCCTGC	ACACGGTTCT	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTTCTGT	TCCAACGACA	4380
	GTTCGCGATT	CTATGCCATG	TTTGCTGCTG	CAGAGAACAA	GGTCCGTGTC	AAGGGCTGAC	4440
	TCTTCCCTGT	TGACGAAGTC	TCTTTTCTTT	AGAGCATTTG	CATTCCCTGC	CTGGGGCGGG	4500
75	CCCTCATCG	CGTCTCTCTA	CCGAAACCTT	GCCTTTCTCG	ATTTTATCTT	TCGCACAGCA	4560
	GTTCGCGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGAGTC	ATATTTTGAT	TATGTATTTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTCTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCTATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTATTATTTA	4800
80	TATTAATAAT	AGCACTGTGC	TAATAACAGT	GCATATTCTT	TTCTATCATT	TTTGTACAGT	4860
	TGCTGTACT	AGACTGTGCG	TTTGTCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCTATTCT	4920
	CTCTAGCTGG	TGGTTTCAAG	GTGCGAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCCC	GCCTCCCCAC	AGCCGCTCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGCGGCTGGA	GACCATGCGG	AGCGCCGTGA	GTCTCAGGCG	CTCCTGCCTT	5100
	CTGTCTGGTT	GTCACTTACT	GTCTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTT	5160
85	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCTCT	TCTTCTTTTT	GCTGTGTTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTCTT	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACCT	5340



GTTGGTTCCA AGCCCTGGAG CCAACTGCTG CTTTGTGAGG TGGCATTCTT TCATTTCCTT 5400  
 ATTCACACAC CTCCACAGTT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTTCCTTT 5460  
 CTCACCCGAG TCGTCGCACA GTCTCTCTCT CTCTCTCCCC TCAAAGTCTG CAACCTTAAAG 5520  
 CAGCTCTTTC TAATCAGTGT CTCACACTGG CGTAGAAGTT TTTGTACTGT AAAGAGACCT 5580  
 ACCTCAGGTT GCTGGTGTCT GTGTGGTTTG GTGTGTTCCC GCAACCCCCC TTTGTGCTGT 5640  
 GGGGCTGGTA GCTCAGGTGG GCGTGGTCAC TGCTGTATC AGTTGAATGG TCAGCGTTGG 5700  
 ATGTGCTGAC CAACTAGACA TTCTGTGCCC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760  
 CAAAAATCTG AAAATGTGAA TAAATATTAT TTGGATTTTG TAAAAAATAA AAAAAAATAA 5820  
 AAAAAAATAA AAAAAAATAA

Seq ID NO: 194 Protein sequence:  
 Protein Accession #: NP\_005679.1

1 11 21 31 41 51  
 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRF TRPLECQDAL ETAARAEGLS 60  
 LDASMHSQLR ILDEEHKPKG YHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSLAR 120  
 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQELNEV GPDAASLRV VWIFCTRILI 180  
 LSIIVCLMITG LAGPSGPAPM VKHLLLEYTQA TESNLQYSL LVLGLLLTRI VRSWSLALTW 240  
 ALNYRTGVRL RGAILTMAPF KILKLNKKE KSLGELINIC SNDGQRMPEA AAVGSLLAGG 300  
 FVVAILMIY NVUILLGPTGF LGSVAVFILFY PAMMPASRLT AYPRKCVAA TDERVQKME 360  
 VLTYIKFIKM YAWVKAFSOS VQKIREERR ILEKAGYFQG ITVGVAPIVV VIASVVTFSV 420  
 HMTLGFDLTA AQAPTIVTVF NSMTFALKVT PFSVKSLSSEA SVAVDRPKSL FLMEEVHMIK 480  
 NKPASPHIKI EMKNATLAWD SSHSIQNSP KLTTPMKKDE RASRGKKEV RQLQTEHOA 540  
 VLAEQKQHLL LBSDERPSPE EEEGKHILG HLRLQRTLES IDLEIQBGKL VGICGSVSGG 600  
 KTSLSAILG QMTLEGSIA ISGTFAVVAQ QAWILNATLR DNILFGKEYD EERYNSVLNS 660  
 CCLRPDLAIL PSSDLTEIGE RANLSGGQR QRLSLARALY SDRSIYILD PLSALDAHVG 720  
 NHIFNSAIRK HLKSKTVLFF THQLQYLVD C DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780  
 FNNLLGETP PVEINSKKET SGSQKKSQDK GPKTGSVKKE KAVKPEBQQL VQLEEKQGS 840  
 VPMVYGVYI QAGGAPLAF VIMALFMLNV GSTAFSTWWL SYWIKQSGN TTVTRGNETS 900  
 VSDSMKDNPH MQYVASYAL SMAVMLILKA IRGVVVFVKT LRASSRLEDE LFRILRSPM 960  
 KFFDTTPTGR ILNRFSKMD EVDVRLPFOA EMFIQNVILV PFCVMIAGV PFVFLVAVGP 1020  
 LVILFVLIH VSRVLIREL RLDNITQSPF LSHITSSIQG LATIHAYNKG QELFHRVQEL 1080  
 LDDNQAPFFL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HQQIPPAYAG LAISYAVOLT 1140  
 GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDNQPE GEVTFENAEM 1200  
 RYRNLPLVL KKVSPFTIKP EKIGIVGRTG SGKSSLGML FRLVELSGGC IKIDGVRISS 1260  
 IGLADLRSL SIIPQEPVLF SGTVRSNLDL FNQYTEDQIW DALERTHME CIAQLPLKLE 1320  
 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIQE TIRAPADCT 1380  
 MLTIAHRLHT VLGSRLIMVL AQGQVVEFDT PSVLLSDSS RPYAMPAAAB NKVAVKG

Seq ID NO: 195 DNA sequence  
 Nucleic Acid Accession #: NM\_006470  
 Coding sequence: 228..1922

1 11 21 31 41 51  
 GCTGCTCTGA GCCTGAGTAC TCTAGCTGCC TTGTGSCCAT GCATCTGGC TGCCATCCAG 60  
 CGCCAGCACA CAGTAATGAG TGGCCGAGCT TCCTCTGGGA GGGAGGAAAC AGTTAAATC 120  
 TTGCAGCAGC TGCAATCATC TAGGCGTGGT TCCTCTGTCT GACTTGGGCT GCACAGATCC 180  
 TGGGCGCAAGG GACAGAGAA AGACAGCCTA GGAGCAGAGC CTCCAGATG GCTGAGTGG 240  
 ATCTAATGGC TCCAGGGCCA CTGCCCAGGG CCACTGCTCA GCCCCAGGCC CCTCTCAGCC 300  
 CAGACTCTGG GTCACCAGC CCAGATTCTG GGTGAGCCAG CCCAGTGGAA GAAGAGGACG 360  
 TGGGCTCTCT GGAGAAGCTT GGCAGGAGA CCGAGGAACA GGACAGCGAC TCTGCAGAGC 420  
 AGGGGATCC TGCTGGTGAG GGGAAAGAGG TCCTGTGTGA CTCTGCCTT GATGACACCA 480  
 GAAGAGTGAA GGCAGTGAAG TCCTGTCTAA CCTGCATGTT GAATTAAGT GAAGAGCACT 540  
 TGCCAGCCGA TCAGGTGAAC ATCAAACTGC AAAGCCACCT GCTGACCGAG CCAAGTGAAG 600  
 ACCACAAGT GCGATACCTG CCTGCCACAC ACAGCCACT GTCTGCTTTC TGCTGCCCTG 660  
 ATCAGCAGTG CATCTGCCAG GACTGTTGCC AGGAGCAGAG TGGCCACAC ATAGTCTCCC 720  
 TGGATGCAGC CCGCAGGAGC AAGGAGGCTG AACTCCAGTG CACCCAGTTA GACTTGGAGC 780  
 GGAACCTCAA GTTGAATGAA AATGCCATCT CCAGGCTCCA GGCTAACCAA AAGTCTGTT 840  
 TGGTGTCTGT GTCAGAGTTC AAAGCGGTGG CTGAAATGCA GTTTGGGGAA CTCCTTGCTG 900  
 CTGTGAGGAA GGCAGGACC AATGTGATGC TCTTCTTAGA GGAGAAGGAG CAAGCTGCGC 960  
 TGAGCCAGGC CAACGATATC AAGGCCACCT TGGAGTACAG GAGTGCCGAG ATGAGAGAAGA 1020  
 GCAAGCAGGA GCTGGAGAGG ATGGCGGCCA TCAGCAACAC TGTCCAGTTC TTGAGGAGT 1080  
 ACTGCAAGTT TAAGAACTAT GAAGACATCA CCTTCCCTAG TGTTAACGTA GGGCTGAAGG 1140  
 ATAACTCTC GGCATCCGCG AAAGTTATCA CGGAATCCAC TGTACACTTA ATCCAGTTGC 1200  
 TGGAGAACTA TAAGAAAAAG CTCACGAGT TTTCCAAGGA AGAGGAGTAT GACATCAGAA 1260  
 CTCAGTGTCT TGCGTGTGTT CAGCGCAAAT ATTGACTTTC CAAACCTGAG CCCAGCACCA 1320  
 GGGAAACAGTT CCTCCAATAT GCGTATGACA TCACGTTTGA CCGGACACA GCACACAAGT 1380  
 ATCTCCGGCT CGAGGAGGAG AACCGCAGG TCACCAACAC CAGGCCCTGG GAGCATCCCT 1440  
 ACCCGGACCT CCCCAGCAGG TTCTGCACT GCGGCGAGT GCTGTCCAG CAGAGTCTGT 1500  
 ACCTGCACAG GTACTATTTT GAGGTGGAGA TCTTGGGGG AGGCACCTAT GTTGGCCTGA 1560  
 CCTGCAAGG CATCGACCGG AAGGGGAGG AGCGCAACAG TTGCATTTC GGAACAAC 1620  
 TCTCTGGAG CCTCCAATGG AACGGGAAGG AGTTACGGC CTGTGACAGT GACATGGAGA 1680  
 CCCCACCTCA AGCTGGCCCT TTCCGGAGG TCAGGCTCTA TATCGACTTC CCGGAGGGA 1740  
 TCCTTCTCT CTATGGCGTA GAGTATGATA CCATGACTCT GGTTCACAG TTTGCTGCA 1800  
 AATTTTCAGA ACAGTCTAT GCTGCCCTCT GGCTTTCCAA GAAGGAAAC GCCATCCGGA 1860  
 TTGTAGATCT GGGAGAGGAA CCGAGAGAGC CAGCACCGTC CTTGGGGGTG ACTGCTCCCT 1920  
 AGACTCCAGG AGCCATATCC CAGACCTTTC CCAGCTACAG TGATGGGATT TGCAATTAG 1980  
 GGTGATTTGT GGGCAGAAAT AACTGTCTGAT GGTAGCTGGC TTTTGAATC CTATGGGGTC 2040  
 TCTGAATGAA AACATTCTCC AGCTGCTCTC TTTTGTCTCA TATGGTCTG TTCTCTATGT 2100  
 GTTTCAGTA ATTCTTTTTT TTTTGTGGA GACGGAGTCT CGCACTGTG CCCAGGCTGG 2160  
 AGAGCAGTGG CGCATCTTG GCTCACTGCA AGCTCCGCT CCGAGTTC AAGCAATCTC 2220  
 CTGCTCAGC CTCGAGTA GCTGGGATTA CAGGTGCTG CCACCAACC CAGCTAATGT 2280  
 TTTGTATTT TAGTAGAGAT GGGGTTTCAC CATGTTGGCC AGGCAGATCT CAACTCTCTG 2340

ACCTCGTGAAT GCACCCCACT CGGCCCTCCCA AAGTGCTGGG ATTACATGCG TGAGCCCACTG 2400  
 CGCCCTGCGCT GTTTGTAGTA ATTTTTAGGC ACCAAATCTC CCTCATCTTC TAGTGCCATT 2460  
 CTCTCTCTGT TTCAAGTAAA TGTCACTAGT TGCCCAGAAT GGATGACCAG GAACCTTAAA 2520  
 GAGTGGCTGA AAAGATGCA GAGTTATCAT AATAAATGTC TAACCTGCGT

Seq ID NO: 196 Protein sequence:  
 Protein Accession #: NP\_006461

1 11 21 31 41 51  
 MAELDLMAPG PLPRATAQPP APLSPDSGSP SPDSGSASPV EEDVGSSEK LGRETEEQDS 60  
 DSAEQGDPAG EGKEVLCDFC LDDTRRVKAV KSLCTCMVNY CEEHLQPHQV NIKLQSHLLT 120  
 EPVVDENWRY CPAHHSPLSA FCCPDQCCIC QDCQEHSGH TIVSLDAARR DKEAELQCTQ 180  
 LDLEKRLKLN ENAISRLQAN QKSVLVSVE VKAVAEMQFG ELLAAVRKAQ ANVMLFLEEK 240  
 EQAALSAQNG IKAHLEYRSA BMEKSKQELE RMAAISNTVQ FLEEYCKPKN TEDITFSPVY 300  
 VGLKDKLSGI RKVITESTVH LIQLLENYKK KLQEFSEKEE YDIRTQVSAV VQRKYWTSKP 360  
 EPSTRBQFLQ YADYITFDPD TAHKYLRLQE ENRKVINTPT WEHPYDLPFS RFLHWRQVLS 420  
 QQSLYLHRY YFVRIFGAGT YVGLTCKGID RKGEERNSCI SGNFWSWLSQ WNGKEFTAWY 480  
 SDMETPLKAG PFRRLGVYID PFGGILSFYG VEYDITMLVH KFAKCFSEPV YAAFWLSKKE 540  
 NAIRIVDLGE EPEKPAFSLG VTAP

Seq ID NO: 197 DNA sequence  
 Nucleic Acid Accession #: NM\_004316  
 Coding sequence: 433-1149

1 11 21 31 41 51  
 CCCGAGACCC GGCACAAGAG AGCGCAGCCT TAGTAGGAGA GGAACGCGAG AC CGCGCAGA 60  
 GCGCGTTCAG CACTGACTTT TGCTGCTGCT TCTGCTTTT TTTTCTTAG AAACAAGAAG 120  
 GCGCCAGCGG CAGCCTCACA CGCGAGCGCC ACGCAGGCT CCGAAGCCA ACCCGCAAG 180  
 GGAGGAGGGG AGGAGGAGG AGGCGCGTG CAGGAGGAG AAAAAGCATT TCACCTTTT 240  
 TTGCTCCAC TCTAAGAAGT CTCCCGGGA TTTGTATAT ATTTTAAAC TTCGTCAGG 300  
 GCTCCCGCTT CATATTTCCT TTCTTTTCCC TCTCTGTTCC TGCAACCAAG TTCTCTCTGT 360  
 GTCCCCCTCG CGGCCCGCGC ACCTCGCGTC CCGGATCGCT CTGATTCGCG GACTCCTTGG 420  
 CCGCCGCTGC GCATGGAAGC CTCTGCCAAG ATGGAGAGCG CGCGCGCGG CCAGCAGCCC 480  
 CAGCCGCGC CGCAGCAGCC CTCTCTGCG CCGCAGCCT GTTCTTTGC CACGCGCGCA 540  
 GCGCGCGCG CGCGAGCGCG CGCAGCGCA GCGCAGAGCG CGCAGCAGCA GCAGCAGCAG 600  
 CAGCAGCAGC AGCAGCAGCA GCAGCGCGCG CAGCTGAGAC CGCGCGCGCA CGGCCAGCCC 660  
 TCAGGGGGG GTACAGAGTC AGCGCCCAAG CAAGTCAAGC GACAGCGCTC GTCTTCGCCC 720  
 GAACCTGATG GCTGCAAAAG CCGGCTCAAC TTCAGCGGCT TTGGCTACAG CCGTCCGCGAG 780  
 CAGCAGCGCG CGCGCTGCG CGCGCGCAAC GAGCGCGAGC GCAACCGCGT CAAGTTGGTC 840  
 AACCTGGGCT TTGCCACCTT TCGGAGCAC GTCCCAACG CGCGCGCAAC CAAGAAGATG 900  
 AGTAAGGTGG AGACACTGCG CTCGCGCGTC GAGTACATCC CGCGCTGCA GCAGCTGCTG 960  
 GAGCAGCATG ACGCGGTGAG CGCGCGCTTC CAGGCAAGCG TCCTGTGCGC CACCATCTCC 1020  
 CCCAATCTCT CCAACGACTT GAATCCATG GCGCGCTCGC CGGTCTCATC CTACTCGTCG 1080  
 GAGCAGGGCT CTACAGACCC GCTCAGCCCC GAGGAGCAGG AGCTTCTCGA CTTCACCAAC 1140  
 TGGTTCGAG GGGCTCGGCC TGGTCAGGCC CTGGTGCAG TGGACTTTGG AAGCAGGGTG 1200  
 ATGCAACAAC CTGCATCTTT AGTCTTTCT TGTCAGTGGC GTTGGGAGGG GGAGAAAAGG 1260  
 AAAAGAAAAA AAAAGAAGAA GAAGAAGAAA AGAGAAGAA AAAAAACGA AACAGTCAA 1320  
 CCAACCCAT CGCAACTTAA CGGAGGCATG CCTGAGAGAG ATGGCTTTCA GAAACGGGA 1380  
 AGCGCTCAGA ACAGTATCTT TGCACTCCAA TCATTCAAG AGATATGAAG AGCAACTGGG 1440  
 ACCTGAGTCA ATGCGCAAAA TGCACTTGT GTGCAAAAGC AGTGGGCTCC TGGCAGAAAG 1500  
 GAGCAGCAGA CGGCTTATAG TAACTCCCAT CACCTCTAAC ACGCACAGCT GAAAGTTCTT 1560  
 GCTCGGGTCC CTTCACCTCC CGCCCTTTT TTAGAGTGCA GTTCTTAGCC CTCTAGAAAC 1620  
 GAGTTGGTGT CTTTC

Seq ID NO: 198 Protein sequence:  
 Protein Accession #: NP\_004307

1 11 21 31 41 51  
 MESSAKMESG GAGQPPQPP QPPFLPPAAC FFATAAAAAA AAAAAAQSA QQQQQQQQQ 60  
 QQQQAPQLRP AADGPPSGGG HKSAPKQVQR QRSSEPELMR CKRRLNFSGF GYSLPQQQPA 120  
 AVARRNERER NRKVLNVLGF ATLREHVPNG AANKKMSKVE TLRSAVEYIR ALQQLLDEHD 180  
 AVSAAFQAGV LSPTISPNYS NDLNSMAGSP VSSYSDEGS YDPLSPPEQE LLDFTNWF

Seq ID NO: 199 DNA sequence  
 Nucleic Acid Accession #: NM\_007015  
 Coding sequence: 1-1005

1 11 21 31 41 51  
 ATGACAGAGA ACTCCGACAA AGTTCCCAT GCCCTGGTGG GACCTGATGA CGTGGAAATC 60  
 TGACGCCCCG CGCGGTACGC TACGCTGACG GTGAAGCCCT CCAGCCCCG CGCGCTGCTC 120  
 AAGGTGGAG CGGTGTCTCT CATTTGGGA CCGTGTGCTG TGCTCTTTGG GCCATCGGG 180  
 GCCTTCTACT TCTGGAAGGG GAGCGACAGT CACATTTACA ATGTCCATTA CACCATGAGT 240  
 ATCAATGGGA AACTACAAGA TGGGTCAATG GAAATAGAGC CTGGGAACAA CTGGAGACC 300  
 TTTAAATGG GAAGTGAAGC TGAAGAAGCA ATTGCAGTTA ATGATTTCGA GAATGGCAATC 360  
 ACAGGAATTC GTTTGCTGCG AGGAGAGAAG TGCTACATTA AAGCGCAAGT GAAGGCTCGT 420  
 ATTCTGAGG TGGGCGCGGT GACCAACAG AGCATCTCCT CCAACTGGA AGGCAAGATC 480  
 ATGCCAGTCA AATATGAAGA AAATTCTCTT ATCTGGGTGG CTGTAGATCA GCCTGTGAAG 540  
 GACACAGCT TCTTAGTTTC TAAGGTGTTA GAACCTGCG GTGACCTTCC TATTTCTGCG 600  
 CTTAAACCAA CTTATCCAAA AGAAATCCAG AGGGAAGAA GAGAAGTGGT AAGAAAAT 660  
 GTTCCAACTA CCACAAAAG ACCACACAGT GGACCAAGGA GCAACCCAGG CGCTGGAAGA 720  
 CTGAATAATG AACCCAGACC CAGTGTTCAG GAGGACTCAC AAGCCTTCAA TCCTGATAAT 780

CCTTATCATC AGCAGGAAGG GGAAAGCATG ACATTGACG CTAGACTGGA TCACGAAGGA 840  
 ATCTGTTGTA TAGAATGTAG GCGGAGCTAC ACCCACTGCC AGAAGATCTG TGAACCCCTG 900  
 GGGGGCTATT ACCCATGGCC TTATAATTAT CAAGGCTGCC GTTGGCCTG CAGAGTCATC 960  
 5 ATGCCATGTA GCTGGTGGGT GCGCCGTATC TTGGGCATGG TGTGAATCA CTTCATATAT 1020  
 CACGTGCTGT AAAATAAGAA CTAGCTGAAG AGACAACCAA AGAAGCATT AAGCAGGTG 1080  
 ATGCTGATGG GACCATAAAA TATTTTACA GCGAGCCTGA GCGTTATTG TTGACACTCT 1140  
 TAACAGAATT TTTTAATCG TTTTCCAGAA CTTTAGTATA TGCAATGCA CTGAAGGGT 1200  
 AGTTCAAGTC TAAATGCCA TAACCCGCTT ATTTGTTATT TTTTATTTC ATTGAATTGC 1260  
 10 CATAGTCTT CCCTTGCTGT CATCTCCAA AGCTATTTCG AAATAAACAC GAAATTTTAC 1320  
 AGTTTGCC

Seq ID NO: 200 Protein sequence:  
 Protein Accession #: NP\_008946

1 11 21 31 41 51  
 MTENS DRVPI ALVGPDDVEF CSPPAYATLT VKPSSPARLL KVGAVVLISG AVLLLFGAIG 60  
 AFYFVKGS DS HYNVHYTMS INGLQDGS M EIDAGNNLET FKMGSAGEEA IAVNDFQNGI 120  
 20 TGIRFAGGEK CYIKAQVKAR IPEVGAUTKQ SISKLEGGKI MPVKYEENSL IWWAVDQPVK 180  
 LNSPLSSKVL ELCDGLPIFW LKPTYPKEIQ RERREVVRI VPTTKRPHS GPRSNPGAGR 240  
 LNNETRPVSQ EDSQAFNPDN FYHQQGESM TFDPRLDHEG ICCIECRRSY THCQKICEPL 300  
 GGYYPWYNY QGCRSACRVI MPCSWVARI LGMV

Seq ID NO: 201 DNA sequence  
 Nucleic Acid Accession #: NM\_000728.2  
 Coding sequence: 112..495

1 11 21 31 41 51  
 30 GTAATAAGAG CGGGGTCTCC GCGGGGAAGG CGCCACAGC AGGTGTGGTG TTCATCCCGG 60  
 GTCGACCGGC CGCTCGCGCT GCCCTGAAAC TCTAGTCGCC AGAGAGGCGG CATGGGTTTC 120  
 CGGAAGTTCT CCCCTTCTCT GGCTCTCAGT ATCTTGGTCC TGTACCAGGC GGGCAGCCTC 180  
 35 CAGGCGGCGC CATTCAAGTC TGCCCTGGAG AGCAGCCAGC ACCCGGCCAC ACTCAGTAAA 240  
 GAGGACGCGC GCCTCCTGCT GGCTGCACCT GTGCAGGACT ATGTGCAGAT GAAGGCCAGT 300  
 GAGCTGAAGC AGGAGCAGGA GACACAGGCG TCCAGCTCCG CTGCCCCAGAA GAGAGCCTGC 360  
 AACACTGCCA CCGTGTGAC TCATCGGCTG GCAGGCTTGC TGAGCAGATC AGGGGGCATG 420  
 GTGAAGACA ACTTGTGACC CACCAATGTG GGTCCAAAG CCTTTGGCAG GCGCGCAGG 480  
 40 GACCTTCAAG CCGAGCAGA TGAATGACTC CAGGAAGAAG GTGTGTCTTA AATCCAATGA 540  
 CATATCCTTA TAAGAGATTC ACTCAGAGA CACATGTGGA GAAGGTGACA TGACAGAGGC 600  
 AAGGAGGCAC AAGCCAAGGA AGTCTGTGTC TACCAGAAGC CAGAATCACA GAACAGTCTC 660  
 TGAAGAAGA GCAGCCCTGC TGACACCTAG AGTTTGGACT TCCAGCTTCC AGAAGTGTGA 720  
 GAGAATAATT TCTGTGTTT TAAGCCACAA AGTTTGTGGT AATTGTGTTAT GACAGCCCTA 780  
 45 GGAAACTAAT ACAATACATT TTCATTATT TTGGGTAAAT GCCTTGGAGT GGGATTGCTG 840  
 GGTATTTTGG AAGAGTGTGA TTTAACTCTG TAAGAAACTG CCAACTATT TTCTGAAGTG 900  
 ACTGTACCAC TTGCGCTTCT TGCCAGCCAC ATATGAGAGC TCTAGTATT CCACAAATAG 960  
 GTATGTAGCA GTATCTCAT CTGCTTTTAA TTGTATTTC CCAATGACT AATGAGCTTG 1020  
 AGCATCTATT TTACCATATG TTTATCACTT TTATTGAAGG GTCTGTTTAA ATCTCTGCT 1080  
 50 AAATTTTGT TGGCTTGCTT GCTTTATTAG TGTGAGTTT TTAGAGCTCT TTATATGTTG 1140  
 TGGATGCAAG ATTGTTTTCA GATATATAGT TTGGAACCTT CCTTCCCTCG AATCTGCGGA 1200  
 TTGCTTTTTC ATTTTCTTAG CAGTGTCTCT CACAGAGAAA AAGTTGTAAT TTGAATAAGA 1260  
 TCCAAATCAT CTTTTTTTTT CTTTATGTA TTGTGCTTT AGTTCATGTC TAAGAACTCT 1320  
 TTGCCTAAT AAGGTCCCAA GGTCAATA ACCTATTCT ATACTTCTT GTAAAAGTTT 1380  
 55 TATAGTTTAT TATTTTATAT GTAGATTAGT GATCTATTT GAGTTAATTT TTGTATAAGG 1440  
 TGAGAGGTGT AGGTGGAAT TCATACCTGT GAATATAGAT ACCCAATTGT TTCAGTGCCA 1500  
 TTTGTAAAA AGACTGTAT TTCAACATT AATTGCCCTT GCACCTTTGT CAAAAAGCAA 1560  
 CTGATCATAT TTGTGTGGGT ATATTCTG GGTCTCAATT CTGCTCATT GATTGATTG 1620  
 60 ACCATCTTT TGCCATGTC ATACTGCCCT GATTAGTGA GTGTAAAGT GAATCTCAAA 1680  
 ACCAGATAAT GTGGGTCTAC CAACATTGTT CATTCTGTT CAAAAGATT TTAGCTACAT 1740  
 CTAAATATT TTCTACATCT TTTATACATT TTAGAATCAG TGTGTTACTA TCTACAAAAT 1800  
 TTTGATGAG ATTTTAAATG GGATTGTGTT AAATCAGTGG GTTAATTTTG GGAGAATTAG 1860  
 CATATTAATA ATATTAAGTC GTTCAATTCA TGAACACAA ACATGTTTTC ACTTATTAG 1920  
 65 GTTTTCTCTG TTTTCTTTT TTTAACAGTG TTCTCAGTT TCAACAGAAA TATTCTACAC 1980  
 ATATCTGTGT AGATTTTAA CTATTTTAT TTTTGTGCT AATGTAAATG GTACTTAAAC 2040  
 ATTTTGTGTT TTAATGTTC ATTGCTAGTA GATAGAAATA CAATATTTAA AATATTAGGA 2100  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 202 Protein sequence:  
 Protein Accession #: NP\_000719.1

1 11 21 31 41 51  
 70 MGRKFPSPFL ALSILVLYQA GSLQAAPFRS ALESSPDAT LSKEDARLLL AALVQDYVQM 60  
 75 KASELKQEQE TQSSSAAQK RACNTATCVT HRLAGLLSRS GGMVKSFPVP TNVGSKAFGR 120  
 RRRDLQA

Seq ID NO: 203 DNA sequence  
 Nucleic Acid Accession #: NM\_001741  
 Coding sequence: 71..496

1 11 21 31 41 51  
 80 CTCTGGCTGG ACGCGCGCGC GCGCGCTGCC ACCGCTCTG ATCCAAGCCA CCTCCCGCCA 60  
 85 GAGAGGTGTC ATGGGCTTCC AAAAGTCTCT CCCCTTCTG GCTCTCAGCA TCTTGTCTCT 120  
 GTTCGAGGCA GGCAGCCTCC ATGCAGCACC ATTCAGTCT GCGCTGGAGA GCAGCCAGC 180  
 AGACCCGGCC ACGCTCAGTG AGGACGAAGC GCGCTCTCT CTGGCTGCAC TGGTGCAGGA 240  
 CTATGTGCAG ATGAAGGCCA GTGAGCTGGA GCAGGAGCAA GAGAGAGAGG GCTCCAGCCT 300

5 GCACAGCCCC AGATCTAAGC GGTGCGGTAA TCTGAGTACT TGCACTGCTGG GCACATACAC 360  
 GCAGGACTTC AACAGTTTC ACACGTTCCC CCAAACTGCA ATTGGGGTGG GAGCACCTGG 420  
 AAAGAAAAGG GATATGTCCA GCGACTTGGG GAGAGACCAT CGCCCTCATG TTAGCATGCC 480  
 CCAGAAATGCC AACTAAACTC CTCCCTTTCC TTCTTAATTT CCCTCTTGCC ATCCTTCCTA 540  
 TAACTTGATG CATGTGGTTT GGTTCCTCTC TGGTGGCTCT TTGGGCTGGT ATTGGTGGCT 600  
 TTCTTTGTGG CAGAGGATGT CTCAAACCTC AGATGGGAGG AAAGAGAGCA GGACTCAGAC 660  
 GTTGGAAAGG AATCACCTGG GAAAATACCA GAAAATGAGG GCGCTTTTGA GTCCCCCAGA 720  
 GATGTCATCA GAGCTCCTCT GTCTGCTTTC TGAATGTGCT GATCATTTGA GGAATAAAAT 780  
 TATTTTCCC C

Seq ID NO: 204 Protein sequence:  
 Protein Accession #: NP\_001732

15 1 11 21 31 41 51  
 MGFKFSPFL ALSILVLLQA GSLHAAPFRS ALESSPADPA TLSEDEARLL LAALVQDYVQ 60  
 MKASELEBQ EREGSSLDSP RSKRCGNLST CMLGTYTQDF NKFTFPQTA IGVGAPGKKR 120  
 DMSSDLERDH RPHVMPQNA N

Seq ID NO: 205 DNA sequence  
 Nucleic Acid Accession #: NM\_005361  
 Coding sequence: 1-945

25 1 11 21 31 41 51  
 ATGCCTCTTG AGCAGAGGAG TCAGCACTGC AAGCCTGAAG AAGGCCTTGA GGCCCCGAGGA 60  
 GAGGCCCTGG CCCTGGTGGG TGCGCAGGCT CCTGCTACTG AGGAGCAGCA GACCGCTTCT 120  
 30 TCCTCTTCTA CTCTAGTGGG AGTTACCCCTG GGGGAGGTGC CTGCTGCCGA CTCACCGAGT 180  
 CCTCCCCACA GTCTCAGGG AGCCTCCAGC TTCTCGACTA CCATCAACTA CACTCTTTGG 240  
 AGACAATCCG ATGAGGGCTC CAGCAACCAA GAAGAGGAGG GGCCAAGAAT GTTTCGGGAC 300  
 CTGGAGTCCG AGTTCCAAGC AGCAATCAGT AGGAAGATGG TTGAGTTGGT TCATTTTCTG 360  
 CTCTCAAGT ATCGAGCCAG GGAGCCGGTC ACAAAGGCAG AATGCTGCTGA GAGTGTCTCT 420  
 35 AGAAATTGCC AGGACTTCTT TCCCGTGATC TTCAGCAAAG CCTCCGAGTA CTTCAGCTG 480  
 GTCTTTGGCA TCGAGTGGT GGAAGTGGTC CCCATCAGCC ACTTGTACAT CCTTGTACCC 540  
 TGCCCTGGGC TCCTCTACGA TGGCCTGCTG GCGCACAATC AGGTCAATGCC CAAGACAGGC 600  
 CTCTGATAA TCGTCTGGC CATAATCGCA ATAGAGGGCG ACTGTGCCCC TGAGGAGAAA 660  
 ATCTGGGAGG AGCTGAGTAT GTTGAGGTG TTTGAGGGGA GGGAGGACAG TGTCTTCGCA 720  
 40 CATCCCAGGA AGCTGCTCAT GCAAGATCTG GTGCAGGAAA ACTACCTGGA GTACCCGCGAG 780  
 GTGCCGGGCA GTGATCCTCG ATGTCTACGAG TTCTGTGGG GTCCAGGGGC CCTCATTGAA 840  
 ACCAGCTATG TGAAGTCTCT GCACCATACA CTAAAGATCG GTGAGAAACC TCACATTTC 900  
 TACCCACCCC TGCATGAACG GGCTTTGAGA GAGGGAGAAG AGTGA

Seq ID NO: 206 Protein sequence:  
 Protein Accession #: NP\_005352

50 1 11 21 31 41 51  
 MPLEQRSQHC KPEEGLEARG EALGLVGAQA PATEEQQTAS SSSTLVEVTL GEVPAADSPS 60  
 PPHSPQGASS FSTTINYTLW RQSDGSSNQ EEBGPRMPFD LESEFQAAS RKMVELVHFL 120  
 LLKYRAREFV TKAEMLSEVL RNCQDFPVI FSKASEYLQL VFGIEVVEV FISHLYILVT 180  
 CLGLSYDGLL GDNQVMPKTG LLIIIVLIIA IECDCAPEEK IWEELSMLLEV FEGREDSVFA 240  
 HPRKLLMQDL VQENYLEYRQ VPGSDPACYE FLWGPRALIE TSYVKVLHHT LKIGGEPHIS 300  
 YPPLHERALR EGEE

Seq ID NO: 207 DNA sequence  
 Nucleic Acid Accession #: NM\_021115  
 Coding sequence: 743-2893

60 1 11 21 31 41 51  
 AAAGGAAGGG AGGGAGGGAG AAAGGAGAAG TTGGTTTAGA GGCCAGCCGG ACAGAGCTTTG 60  
 GGCAACGCCC TTAGGAGGGC CACCTCAGA GTCTGACAGC AGGTGAAGGT CCTAAATCTC 120  
 65 CCCAACTAA CTGGTGTCTT TTCTCTCTT CCAAGATGCT CTTCGCGAGG GAGATGCTAG 180  
 CCCTTTGGGT CTTACCTCC TGCCCTCAGG AGCCCGGAG AGAGGCAGTC CTGGCAAAGA 240  
 GCACCCGTAA GAGAGAGTGG TAACAGCGCC CCCAGTTCC TCACAGTCGG CGGAAGTGCT 300  
 GGGCGAGCTG GTGCTGGATG GGACCGCACC CTCTGCACAT CAGACATCC CAGCCCTGTC 360  
 ACCGCTGCTT CCAGAGGAGG CCGCCCCCAA GCACGCTTGG CCCCCAAGA AGAAACTGCC 420  
 70 TTGCTCAAG CAGGTGAAT CTGCCAGGAA GCAGCTGAGG CCCAAGGCCA CCTCCGCGAGC 480  
 CACTGTCCAA AGGGCAGGGT CCCAGCCAGC GTCCAGGGC CTAGATCTCC TCTCTCTCTC 540  
 CACGGAGAAG CTGGGCCAC CGGGGGACCC GGACCCCATC GTGGCTCCG AGGAGGCATC 600  
 AGAAGTGCCC CTTTGGCTGG ACGAAAGGA GAGTGGGGT CCTACAACAC CCGCACCCCT 660  
 GCAAATCTCC CCCTTCACIT CGCAGCCCTA TGTGCCCCAC AACTCCCCC AGAGGCCAGA 720  
 75 ACCCGGGGAG CTTGGGCTTG ACATGGCCCA GGAGGCCCCC CAGGAGGACA CAGCCCCAT 780  
 GGCCCTGATG GACAAAGGTG AGAATGAGCT GACTGGGTCA GCCTCAGAGG AGAGCCAGGA 840  
 GACCACTACC TCACCATTA TCACCAACC GGTATCACC ACOGAGCAGG CACCAGCTCT 900  
 CTGCACTGTG AGCTTCTCCA ATCTGAGGG GTACATTGAC TCCAGCGACT ACCCACTGCT 960  
 GCGCCCTAAC AACCTTTCTGG AGTGACATA CAACGTGACA GTCTACACTG GCTATGGGGT 1020  
 80 GGAGCTCCAG GTGAAGAGTG TGAACCTGTC CGATGGGGAA CTGCTCTCCA TCGCGGGGGT 1080  
 GGAAGGCCCT ACCAACACCA TCTCGGCTA CTTCGGGACC TTCCAGGACG ACGGCCTTGG 1200  
 GACCTTCCAG CTTCACTACC AGGCCTTCAT GCTGAGCTGC AACTTTCCCC GCGGCTGGA 1260  
 CTCTGGGGAT GTACAGGTGA TGGACCTGCA CTCAGGTGGG GTGGCCCACT TTCAGTGCCA 1320  
 85 CTTGGGCTAT GAGCTCCAGG GCGCTAAGAT GCTGACATGC ATCAATGCC CCAAGCCGCA 1380  
 CTGGAGCAGC CTGGAGCCCA TCTGCTCAGC TCCTTTGTGA GGGGCACTGC ACAATGCCAC 1440  
 CATCGGCCGC GTCTCTCCC CAAGTTACCC TGAACAACA AATGGGAGCC AATTCTGCAT 1500  
 CTGGAGGATT GAAGCTCCAG AGGCCAGAA CTTGCACCTG CACTTTGAGA GGCTGTTGCT 1560

WO 02/086443

PCT/US02/12476

5  
10  
15  
20  
25  
30

GCATGACAAAG GACAGGATGA OGGTTACAG OGGGCAGACC AACAAGTCAG CTCTTCTCTA 1620  
CGACTCCCTT CAAACCGAGA GTGTCCCTTT TGAGGGCCTG CTGAGCGAAG GCAACACCAT 1680  
CCGCAATCGAG TTCAGTCCG ACCAGGCCCG GCGGCCCTCC ACCTTCAACA TCCGATTGTA 1740  
AGCGTTTGAG AAAGGCCACT GCTATGAGCC CTACATCCAG AATGGGAAGT TCATACATC 1800  
CGACCCGACC TATAACATTG GGAATATAGT GGAGTTCACC TGCAGCCCGG GCCACTCCCT 1860  
GGAGCAGGGC CCGGCCATCA TCGAATGCAT CAATGTGCGG GACCCATACT GGAATGACAC 1920  
AGAGCCCTCG TGCAGAGCCA TGTTGGTGG GGAGCTCTCT GCTGTGGCTG GGGTGGTATT 1980  
GTCCCCCAAC TGGCCCGAGC CCTACGTGGA AGGTGAAGAT TGTATCTGGA AGATCCAGCT 2040  
GGGAGAAGAG AAACCGATCT TCTTAGATAT CCAGTTCCTG AATCTGAGCA ACAGTGACAT 2100  
CTTGACCATC TAGCATGGCG ACAGGTCAT GCCCCACATC TTGGGGCAGT ACCTTGGGAA 2160  
CAGTGGCCCC CAGAAACTGT ACTCTCCAC GCCAGACTTA ACCATCCAGT TCCATTCCGA 2220  
CCCTGCTGGC CTCTCTTTG GAAAGGGCCA GGGATTATC ATGAACCTACA TAGAGGTATC 2280  
AAGGAATGAG TCCTGCTCGG ATTTACCCGA GATCCAGAAT GGCTGGAAAA CCACTTCTCA 2340  
CACGAGTTG GTGCGGGGAG CCAGAATCAC CTACAGTGT GACCCCGGCT ATGACATCGT 2400  
GGGAGTGAC ACCCTCACCT GCCAGTGGGA CCTCAGCTGG AGCAGCGACC CCCCATTG 2460  
TGAGAAATTT ATGTATCTGA CCGACCCCGG AGAGGTGGAT CACTGAGACC GCTTAATTTT 2520  
GGATCCTGTG CTGCTGGTGG GAGCAACCAT CCAATACACC TGCAACCCCG GTTTTGTGCT 2580  
TGAAAGGAGT TCTCTTTCTG CCTGTACAG CCGTGAAACA GGGACTCCCA TCTGGACGTC 2640  
TGCCTGCCCC CACTGGGTTT CAGAAGCGGC AGCAGAGACG TCGCTGGAAG GGGGGAACAT 2700  
GGCCCTGGCT ATCTTCATCC CGGTCTCTAT CATCTCCTTA CTGCTGGGAG GAGCCTACAT 2760  
TTACATCACA AGATGTGCGT ACTATTCCAA CCTCGCCTG CCTCTGATG ACTCCCACCC 2820  
CTACAGCCAG ATCACCCTGG AAACCGAGTT TGACAAACCC ATTTACGAGA CAGGGGGAAC 2880  
CCAAAGGTT TAGGGTTTCA TTTAAAAAGA GGTACCCCTT AAAAAGGGGC TTGTGAACTC 2940  
AAGCCCAATT TCCCGGAGAC ATTTATCCAA AGGCCCTGGG GGCCTTGATT TAAACCCCA 3000  
AAAGGCGGCT GTTTTGTGTT TAACTTTT AACAAAGGT TACGGGTTT TTCCCGGAT 3060  
TTTATAAATT TTAAAGTG

Seq ID NO: 208 Protein sequence:  
Protein Accession #: NP\_066938

35  
40  
45

1 11 21 31 41 51  
MAQEAPQEDT SPMALMDKGE NELTGSASEE SQETTISTII TITVITTEQA PALCSVSFSN 60  
PEGYIDSSDY PLLPLANPLE CTYNVTVTG YGVELQVKS NLSGGLLSI RGVDPGLTV 120  
LANQTLLEVG QVIRSPINTI SVYFRTPQDD GLGTFQLHYQ AFMLSCNFFR RPDSDGVITM 180  
DLHSGGVAHF HCHLGYELQG AKMLTCINAS KPHWSSQEPF CSAPOGGAVH NATIGRVLSP 240  
SYFENINGSQ FCTCTIEAPE GQKLHLHFER LLLKDKDRMT VHSQGTNKA LLYDSLQTES 300  
VPFEGLLSEG NTRIEFTSD QARAASFTPI RFEAFKGGHC YEPYIQNGNF TTSDPYINIG 360  
TIVPEPTCDG HSLBQGPAIL ECINVRDPYV NDTEPLCRM CGSELSAVAG VVLSFNWPEP 420  
YVBOEDCIWK IHVGEKRIFF LDIQFLNLSN SDILTIYDGD EVMPHILGOY LGNSGPKRLY 480  
SSTPDLTIQF HSDPAGLIFG KGGFIMNYI EVSRNDSQSD LPEIQNGWKT TSHTELVRGA 540  
RITYQCDPGY DIVGSDTLTC QWDLWSWSDP PFCEKIMYCT DPGEVDHSTR LISDPVLLVG 600  
TTIQTCTNPG FVLEGGSSLLT CYSRBTGTPI WTSRLFCVSV EAAAEYSLEG GNMALAFIP 660  
VLIISLLGG AYIYITRCRY YSNLRLPLMY SHFYQITVE TEFDNPIYET GGTQKRV

Seq ID NO: 209 DNA sequence  
Nucleic Acid Accession #: NM\_001327.1  
Coding sequence: 89-631

50  
55  
60  
65

1 11 21 31 41 51  
AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60  
CTGAGAGCCG GGCAGAGGCT CCGAGGCCAT GCAGGCCGAA GGCAGGGGCA CAGGGGGTTC 120  
GACGGGCGAT GCTGATGGCC CAGGAGGCCG TGCCATTCTT GATGGCCAG GGGGCAATGC 180  
TGGCGGCCCA GGAGAGGCGG GTGCCACGGG CGGCAGAGGT CCGCGGGGCG CAGGGGCAAGC 240  
AAGGGCCTCG GGGCCGGGAG GAGGCGGCCG CGGGGTCCG CATGGGGCG CGGCTTCAGG 300  
GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGAG AGCCGCTGCT TTGAGTCTTA 360  
CCTCGCCATG CCTTTCGCGA CACCCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420  
GGATGCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TGTCCGGCAA 480  
CATACTGACT ATCCGACTGA CTGCTGCAGA CCACGCCCAA CTGCAGCTCT CCATCAGCTC 540  
CTGTCTCCAG CAGCTTTCCC TGTGATGTG GATCACGCG GATCTTCTGC CCGTGTTTTT 600  
GGCTCAGCCT CCTCAGGGC AGAGGCGCTA AGCCAGCCT GCGGCCCTT CTTAGGTCTAT 660  
GCCTCTCCCT CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720  
GTTTGTCTCT GGAGGAGGAC GGCCTACATG TTTGTTCTG TAGAAAAATA AACTGAGCTA

Seq ID NO: 210 Protein sequence:  
Protein Accession #: NP\_001318.1

70  
75

1 11 21 31 41 51  
MQAEGRTGG STGDADGPGG PGIPDGPGEN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60  
PRGFHGGAA GNGCCRCGA RGPESRLLEF YLAMPFATPM EAEALARRSLA QDAPPLFVPG 120  
VLLKFTVSG NLTITRLTAA DHRQLQLSIS SCLQLSLLM WITQCFPLVF LAQPPSGQRR

Seq ID NO: 211 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 52-459

80  
85

1 11 21 31 41 51  
CCTCGTGGGC CCTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGGAGC CATGCAGGCC 60  
GAAGGCCAGG GCACAGGGGG TTGACGGGC GATGCTGATG GCCAGGAGG CCTTGGCATT 120  
CCTGATGGCC CAGGGGGCAA TGCTGGCGGC CCAGGAGAGG CGGCTGCCAC GGGCGGCCGA 180  
GGTCCCGGGG GCGCAGGGGC AGCAAGGGCC TCGGGCCGA GAGGAGGGCG CCGCGGGGGT 240  
CGCATGGCG GTCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTCGGGGC CAGGAGGCCG 300

## WO 02/086443

PCT/US02/12476

5 GACAGCGGCC TGCCTCAGTT CCGACTGACT GCTGCAGACC ACCGCCAACT GCAGCTCTCC 360  
ATCAGCTCCT GTCTCCAGCA GCTTCCCTG TTGATGTGGA TCACGCAGTG CTTTCTGCC 420  
GTGTTTTTGG CTCAGGCTCC CTCAGGGCAG AGGGGCTAAG CCCAGCTCG CGCCCCCTCC 480  
TAGGTCTATG CTCTCTCCCT AGGGAATGGT CCCAGCACGA GTGGCCAGTT CATTGTGGGG 540  
GCCTGATTGT TTGTCTGTGG AGGAGGACGG CTACATGTT TGTTCTGTGA GAAATAAAG 600  
CTGAGCTA

Seq ID NO: 212 Protein sequence:  
Protein Accession #: Bos sequence

10 1 11 21 31 41 51  
MQAEGQGTGG STGDADGPGG PGIPDGPGGN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60  
PRGPHGGAAS AQDGRCPGCA RRPDSRLQLF RLTAADHRQL QLSISSCLQQ LSLMNIQTQC 120  
15 PLPVFLAQAP SQQRR

Seq ID NO: 213 DNA sequence  
Nucleic Acid Accession #: NM\_000555  
Coding sequence: 416..1498

20 1 11 21 31 41 51  
CTTATTTTTT ATGAATGTGC GATAGCTGCA CCAGCTTGGT GGGGAAGGG TTTGATGAAT 60  
AGCACAAAGA CACTGGCTGT TCCCTGGAGG CTGTCCCTTT AAAGGAGAAT CTTAGTTTAT 120  
25 TCTGGGGGGA GGGGATGCAC ACATTAGAGT AGGAAAGAGG GCTTGAATAA AAATGAAAAAC 180  
ACTCCCCCTT CATAGTCATT GTACTGAAAT GCAAAGACTG CTTCTTAAGC TGGAGATGCT 240  
AACCTTGGGT AGCTCCCTCT GTTCTCTTCA AGGGGAATTT TGTGAGGCTA TGGATTCTAT 300  
TACAACTGTT AGTCATGTGG GCATGTGTGA GGAACAGAT GCCAGTTTGA ATGTATTTAG 360  
CCCGAAGTTC CAATTTGATA GGAGCCACTG TCAGTCTCTG AGGTTCCACC AAAATATGGA 420  
30 ACTTGATTTT GGACACTTTG ACGAAAGAGA TAAGACATCC AGGAACATGC GAGGCTCCCG 480  
GATGAATGGG TTGCTAGGCC CCACTCACAG CGCCCACTGT AGCTTCTACC GAACAGAAAC 540  
CTTGAGGCA CTGAGTAATG AGAAGAAAGC CAAGAAGGTA CGTTTCTACC GCAATGGGGA 600  
CGCTACTTTC AAGGGGATTT TGTAGCTGT GTCTCTGAC CGTTTTCGCA GCTTTGACGC 660  
CTTGCTGGCT GACCTGACGC GATCTCTGTC TGACACATC AACCTGCTCT AGGGAGTGGC 720  
35 TTACATTTC ACCATTGATG GATCCAGAA GATCGGAAGC ATGGATGAAC TGGAGGAAGG 780  
GGAAAGCTAT GTCTGTCTCT CAGACACTT CTTTAAAAAG GTGGAGTACA CCAAGAAATGT 840  
CAATCCCAAC TGGTCTGTCA ACGTAAAAAC ATCTGCCAAT ATGAAAGCCC CCCAGTCTCT 900  
GGCTAGCAGC AACAGTGCAC AGGCCAGGGA GAACAAGGAC TTTGTGCGCC CCAAGCTGGT 960  
TACCATCATC CGCAGTGGGG TGAAGCCTCG GAAGGCTGTG CGTGTGCTTC TGAACAGAA 1020  
40 GACAGCCCACT TCTTTTGAGC AAGTCCCTAC TGATATCACA GAAGCCATCA AACTGGAGAC 1080  
CGGGGTGTGC AAAAAGCTCT ACACCTGGA TGGAAGACAG GTAACCTTGT TCCATGATT 1140  
CTTTGGTGAT GATGATGTGT TTATGCTGCT TGGTCTGAA AAATTCGCT ATGCTCAGGA 1200  
TGATTTTTCT CTGATATGCG ATGAATGCG AGTCATGAAG GGAAACCCAT CAGCCACAGC 1260  
45 TGGCCCAAG GCATCCCAAC CACTCAGAA GACTTCAGCC AAGAGCCCTG GTCCTATGCG 1320  
CCGAAGCAAG TCTCCAGCT ACTCAGCAA CGGAACCTCC AGCAGCCAGC TCTCTACCCC 1380  
CAAGTCTAAG CAGTCTCCCA TCTCTAGGCC CACCACTCTT GGCAGCTCC GGAAGCACAA 1440  
GGACCTGTAC CTGCTCTGT CTTGGAGTGA CTCGGACTCG CTTGGTGATT CCATGTAAAG 1500  
GAGGGGAGAG TGTCCAGAG CCAGAGTACA AATCCAAAGC TATCATGTGA GTAGGGTACT 1560  
50 TCTGCTCAAG TGTCCACAG GGTCTATGGT GCTTTCAAGT TTTTATTTTG TTGTTGTGTG 1620  
TATTTTGAAG AACACATCTT AATATGTTGG GTTTATTTTC CTGTGATTTT TCTCTGGGC 1680  
CACTGATCCA CAGTTACCAA TTATGAGAGA TAGATTGATA ACCATCTTTT GGGCAGCAT 1740  
TCCAGGGATG CAAATGTGTC TAGTCCATGA CTTTCAATG GAAAGCTTAG GGGCCTGGGG 1800  
TAAATTTGCC CCGTTTAAAT TTGCCAAAC AGTTTTCTTT TTGTAGAGGG GTGTTTAAAT 1860  
ATACAGCAAT TAAAGAGTTT GTGTGGGGA AAAAAAACT CATTGGCAGA TCCAAGATG 1920  
55 ACAAACCAA GTGCCCTTT TCTCTGATC TCAAGATGG TGGAGGACCC TGAAGGACA 1980  
GCAAGGCAGC TCCCAGCCT CACTCTTCAC TCTGATTGA GGGCCGGGT TGTGTCCAG 2040  
CACCAATTTT GGTCTCAAT GGGGAGAAAT AAACCAACAA CTTAATATTG TGACACCAGA 2100  
TGCTTAGGAT TCTGGTCTG GTTAGCTAA GAGAATAGAC AGAATTGGAA AATACTGCAG 2160  
ACATTTCCGA AGAGTTTATA AAGCACAGTG AATTCCTGGT CAATCTCTCC ACTGAGGCAA 2220  
60 TTTGGAATCA ATAGCAATG GATAATAGTT TGGAGTAAGG GACTTCATAT ACCTGATTC 2280  
TCTAGAAGGC TGTCTAATAT ACCACATGAT TACATGAAT GTATGATATC CATCTATCT 2340  
TGTTCTATTG AATGCCCTGT TAACAGCCAA CACTGAAAC ACTGTGAGAA TTTGTTTTCA 2400  
GGTCTGACAC CTTTCACTT CTTTATATAG CAAGAAATCA ATATCCTTTT TATAAAATTT 2460  
CATGTCTGTA TTTGAGGAGC AAATCTTTCA GGTCTCTTTT TTATAAACTG GTGATTTTTC 2520  
65 TTTTGTCTAA AAAACACATG AAGAAAATTT ACCAGAAAAA AAAAAAAAG CCGAAGATA 2580  
ATGTTATTTA GAAATATGTC TGTCACTGCC AAACAGTAAC CTCAGGAGA AAACAGATG 2640  
AATAGCAGAG GCCAATTCAA TAGAATCAGT TTTTGTATG CTTTAAACA GTTATGCTTG 2700  
CATTAATAAT TTCATGTGG ACCAGACATT CTAATTATAT TTTAAATGAA ATGTTACAGC 2760  
ATATTTTAA GAACTCTTTT TATCTATAAT CTAATATTTT CATACTGAAG ACACAGAAAT 2820  
70 CTTTCACTTG CTTTAAATAT TAGAAAGGAT TTCTCTTAC TAAGGACTGA TCATTTGAAA 2880  
TAGTTTTCAG TCTTTTGAGA TACAGTTTAA TAACACTGCT TTTTTTTCC TGTAAACATA 2940  
GCCATAATG GCAAAAACAA CTAATTTTAA TTGAAGGTCT TGCTTGCCAN TCCTGTGTTG 3000  
GCTTTNACCA AATATAAAAA TTCCCTTATT CCTTGGTAAT GGTGCAATN TTTGAAAGG 3060  
75 CACAGCATCC AAACCAAGCT GCTGTTTGGC TACTGAATGG CTGCACTGTG TTCCTCCACT 3120  
CTAAATGGAA TGAGCTTGCT GTGTGTGTGT GTGTGTGTGT TGGGAGGGGG TGGTGATGT 3180  
GTGTGTGTGT GTGTGATCT CGAGCTGCTT CAAATTAAG AAATACTACA AGACACCCCT 3240  
GTAATGGATT GGTGGCACT GGTGGCACT GCTGATGTGC ACTGTGTAGG GGGGAACCCA 3300  
GTGTGTGTGT GGTATCTCAA ATGCCCTAG ACAAGCTTCA GATGTCTGTA GCTACCAAAA 3360  
ACATTTTGG TTAGAGAAA GTGAGATGAT GGTAGTACTG GTTCTGGTG AAATGAAAA 3420  
80 ACCCAAAATG ATGAGGATCT CTTTGTGCC CTTCTCTTTT TTTTGTAAAC CCATTCAAA 3480  
CCATTAATAA GCCCATTTTA CTAANCCCTT ATTTCTTCT AGAAGCTCAG GGTTNTCTTA 3540  
GTGCTCCCA NAACATTTT TAGTTAATG GGAAGAGTG ATACTTGGAT TAGGGGGTGT 3600  
GGGCATAAAG AATGGTGGGA GGCCTGATT TAAATTCAG GCCAGAACCC CCAATGACTC 3660  
85 CACCCATAGT NTCATTTAG GTTCATTTA GTCCATCACC TTTATTTTAA GTTGAGGAG 3720  
TGGAGGCTG TAGAGACAG GACCAGGGA AGAATCCAGA TTTCTTATG CTGTGGCTTC 3780  
ACACTAGCTC TGTAGATATT TCTTGTATT CGGTATATGT ACTACTAGAA AATACCAAT 3840  
GGATATATT TCTTAGGAT AACCTTTGAA CCAACAAAT TCAATAACAA TAGTACATCT 3900

	TCCATCTTAC	TTTTAAATCGA	GTATAAGGAA	ATGTTTCTTT	ATGGCCATT	TGGAGGGAGC	3960
	AGGGGATGAG	GCTTGGCATA	GTCCAAAATT	TAAGNCTCCA	ATAATTAATT	GCATTTTAAA	4020
	TTGTTTTAAA	TTGGCCCACT	TTCAAGGCAA	TTTTTTTGT	GTGTCTGTAA	CTGAGCTCCT	4080
5	CCACCCCTGT	CATTCACTTC	CAATTTTACC	CAATCCAATT	TTAGCACTCA	AGTTCCTATT	4140
	TGTTAAATTT	TGCAAGGTCT	ACACACATCA	AGTCAGCAAG	CATTTGCCAC	CACCTCCCTAT	4200
	ACTTCTCCCT	CTTTTATACA	CACACACACA	CACACACACA	CACAATCCAT	CTCTTGCTTG	4260
	TTCTTACCTC	CCTGATTTTT	CTTCCCTACA	GAAATAGAAA	TAGGGACAAA	GAAGGGGAAA	4320
	ATGTATATAT	TGGGCTGGG	CTGAACAAC	AACCTTCATA	GTAGTATTAA	CTAGGGGTAA	4380
10	ATTGAGAGAA	AAGCTCCTTT	TCTCTTCACT	GTTTTGGAAA	GGATAGCCAT	TAGCATGACT	4440
	GCTTTGTGTC	CTTATGGACT	TTAGTATTAG	CCTAGATTGA	ATTATAGCGT	TTTTCTAGCT	4500
	GAAGGAACCT	TAAGATCACA	TCATCTACTC	CTCTACTCCA	AATTTCTCAT	TCTTCAGGCC	4560
	AGGAAACCGA	GACACAGAGG	TAAAGTAATT	TCCCAAGGT	CACACAGCTG	GCTGGGGCAG	4620
	GATTGGGTTT	ACAACCCACA	TCTCCTGGCT	CTTATTCAG	GGCCTTTTCC	CACCTAAGTAG	4680
15	TATTGCCCTC	CATTAGGCTC	CTGAGAGTTA	TTTCTCAGGG	TCATGTTGCA	TCTTGGAGCC	4740
	ACATGCTGCT	GCCTGATCTC	CAGTGGGAAA	TNCACCCAGC	AACCTAATAC	AGCCCCCTTT	4800
	CCCTGCAATC	ACCTGGTTCC	CATCCACATG	GGTTGCAGAT	GTCTTGAAG	AGAGTGAGGC	4860
	ATTGAGGGCC	AATAGGAGCA	ATGGGGTCCC	TGGCCTTGTC	CATCTGATTC	AGGAGATCAC	4920
	TGCTCCATCG	TGAGGAGCCC	TCTGAATAGC	CCCCACTGA	ATGCTTGCC	TGCCCAATG	4980
20	GAATGGAGGA	AGATTGATTT	TCTCCATCAG	TTCACCTGT	GTCTCTCAT	AATGGTTGGT	5040
	CTTTCCAGCG	TGAGGGAATT	GTTCCTTGT	TCCANAGTAN	AAAAAAGAAA	GAGTGGAACA	5100
	ATANCTTTGT	TGATCTTAAC	TTTCTGAGAT	GGCTTTTCAA	CATTAAAAAA	AAACTAGTGT	5160
	GGTACCATTG	ACTGGCANGA	TTTNTTTTAG	AATATGGGAG	TAAGATGAGG	TAGAGAAAAAT	5220
	AACCTGGTCT	CACGTGGTGT	GCCTCATCC	ACAATGTCCC	CAAGCCATC	CTGCTNTGAT	5280
25	GAGGACAATT	TCCAGGTATA	AGCAAGGGGC	TTTGTGACAA	AAATGTACCC	TGGCTGATGT	5340
	TAAACATTGG	CTCCTGTGTT	TGCACCAAAA	TAGCAAGCTG	TGTGCTCTAT	ACACTCTTCC	5400
	CATCGTCTTG	TGTACTCTGC	TCCCTGGGCC	TTCCACAGCA	GAAACCCAGG	CAAAAGGGTC	5460
	CAACACATG	GTTCCTCTG	CTGCAAGGCT	NTTCTGGGA	ACTAAGGGGG	TATTTATTAG	5520
	TTCACTGNTA	AGAGACTGCT	TTCTGGGCTT	ACCCACTCC	TCAGTACTT	CTCTCTCCTT	5580
30	CCTCCTTCTC	CTCCACAGTC	ACAAGTAACC	AAGGAACCTG	AAAGTGGATG	TGTAGCTATT	5640
	TGAGGAGGCG	AAGGAACCTC	GAGATTCTTC	TTTGAATCCT	TTAGTCCAAG	TCTTAGACCA	5700
	GTGATTGGTG	CTTACCTTGA	ACAAAATTTT	GTCTGTGTTT	CTAATCCTT	CAATCTNTG	5760
	GGTACAATGC	TCCCAATCAC	CCTGCACATT	TGATTTAAAA	TGGCTTTTAT	TTTTTAAAAA	5820
	TCCATATCCC	TAGGACAAGA	NAACAGGATG	CCTATATCCC	CAAAATGAGC	TCCAGGACAC	5880
35	TGATGGGAAT	GATCCCAANG	ATCAGCCAC	CTCAGAAAAC	GTCTGTGCCA	ANAGACTTCC	5940
	CCAGATAGAA	NCACTGGGAG	AGTGGTTTGA	ACGACTTCTT	TTATGGTTGT	CCAGTTTGCT	6000
	ATGGAAATAA	AAGGCATTGA	TTTTTTAAAA	AAGATGATTG	GAACCTGTCT	TTGGCCACAT	6060
	AGGGCCACTT	GGATCCATT	CCAGGCTTAA	CTCATATATT	GCCTTCAGTG	AAGGGCTTTG	6120
	GCTTTAAGTC	CCAGACTGGT	CTCCCAAGTG	AACCATAGGT	GTITTTGGAGC	TCATCTGGGG	6180
40	TGAGGCATGA	GAATGTTGCC	CCATCTATCC	CTTCAGGAAA	AGGTGCCCTC	CCTCCCTTTC	6240
	TCCTAAAGCC	TGGTCCCCAA	AAATTGTTTT	TGTCCTCAAA	AGTCTAGTAT	GGTCTTTATA	6300
	CACCANACT	CTTAGTGTGG	CGTCTGCTCT	TGTTTCTCTG	TTAAGGATCT	ATGCANACCT	6360
	CCCGCTTTGG	CTTAGCTAGC	GTGACATTGG	CTATCATTTG	ACAAGACTAA	CTTTTTTTTT	6420
45	TTTTTTTTTG	ACTGAGTCTC	CCTCTGTAC	CTAGGCTGGA	GTGCACTGGC	ACAATCTTGG	6480
	CTGCTGCAAA	CCTTCAACCT	TCACTCCCCA	GGTGAAGCG	ATTCCTCTGC	CTCAGTCTCC	6540
	CGAGTAGCTG	GGATTACAGG	CGTGGGCCAC	CAAACTCTGC	TATTTTTTTA	TTATATTAT	6600
	TTTTAGTAGA	GATGGGGTTT	CACCATGTTG	GCCAGACTGG	TCTTGAACCT	TTGGCTCAA	6660
	ATTATCTGCC	CACCTCGGCC	TCCCAAGTGG	CTGGGATTAC	AGGCATGAGC	ACCATGCCCA	6720
	GCTGACAGGA	CTAATTTTTT	ATCCCTTGGT	TTATGGCTT	CAACATCTTC	TGGAACTAGA	6780
50	GGTGATTTTT	TCTTACCTTG	GATGCTGAG	ACTAGGGGAG	TATAGAAATC	CAATGGTAA	6840
	TTAAGGCATC	TTTCTGCTCC	TGATCAGAAG	GGCAGGTTAG	TTGGGAGAGG	TCAGATGGCA	6900
	CAACAGAAGT	CACCTTGTA	GTAAGGCAAA	GACTTTGAAG	GCATTAGCGT	TTCTCATTAC	6960
	TTAGGTCAAT	AACTTGAGG	GAATCAATGG	CTTTTTTGCC	GCTCTACCTC	TTTGTGTATC	7020
	TCCTTGACTT	TTCTTCTCT	GTCTAGTTTC	CTCTGTTCTC	AGTTTATATT	CTATGTTATC	7080
55	AGTCTCTCTT	TCCACAGTAC	AAACATCCAT	CCCTTCTCCT	GTGCAATTCT	GTCTCTCCTC	7140
	CTTATATATC	TTATTGTATC	TTTTTCTTTC	CTCCTGTCT	AGGCATTGGG	CACTGCTCTC	7200
	TTCTTAGCCT	GTGATTTTGC	CTTGGGACTG	ATGATAAATT	ATTTCCAGAT	TCAATCAGCC	7260
	CTGGTCTTAC	CCAGTCCCAA	TCAGAAGTAT	GTGGTGGGG	AATCAACCTG	ATCCTGGCCC	7320
	TTTCTTCTTC	TTCATTTTCA	TTGTAATCC	CCCTCAGCAG	ATCTTTACAA	GCAGTTTCTT	7380
60	TATAGCTCAT	GTATCTTTAG	GTCTTTGGCT	TCCAAGCACT	GTACAGAATA	CTTTGTGTTT	7440
	CCTTTTAGGT	CTGACATTTT	GTGGAGCAGT	GAAGCGTGCT	CAGAGACATA	ATCAGCTGAA	7500
	GAGAAAAAAT	CCACCCATGG	ATTTATATCA	GCTAAATACT	AATAATTGAT	TTTGTTTGAT	7560
	GTGCCCATAA	TTTTAAAGC	TGCAATATAA	TATAATGAGG	GACCACAGGT	AATTTCTCCT	7620
	GTCAATTTGT	TTGGCTGGAT	GGGGGTGGGG	GAGTAATTGC	TTAAAGTTT	ACCATTAAC	7680
65	ATTAAACTCT	CTATAATAAT	CTTGTTTGGG	GCTTGCTAAC	TGTTGAGCTG	TTTAACTAA	7740
	ACTGGTAGGC	AATCGGAGTT	GATTTAAATG	AAAAGATAAT	TTAACAAATC	TATACTATAA	7800
	AAAGAGACAT	TTGCTTAATT	GACATGTATT	TTTTCTTCT	GAGTCACCTA	AACATTACT	7860
	CTTGACACCA	ACTGTTCTAG	ATACTGAATA	GACAGTCCAT	ATAAGAGAAA	TTAGTGGACC	7920
	TAAAGAAGCC	AGATTGTAGG	TGTTAATTTA	TTAAACAGAA	TTGCAAGGCC	CTTGGAAATG	7980
70	TCACTGCTTG	GCAATACCAT	ATGGCATGCC	AAAATTTACA	ATGACTTTTC	TTTATAAGTT	8040
	ATCCAAAGGG	GATTTGAACA	AGTAAGAGGT	TATGCCAAAA	TGCTCTCAAT	GTATGGTCTC	8100
	GTAATATATT	GCAGCTTGAA	GCCAATGATC	CCTTATGACT	TGTATACAA	TATATCATGT	8160
	TTTATTGAAT	TTTGCAATTC	CCAAGTGTGG	TAAGTCTTTA	AAATGTTTTT	GATCACCTTT	8220
	NTGTGCCATT	AAACTGTAC	AGAAAATGTT	TTTATGGCCA	TTTTCAAAGG	GAGAAAGTTT	8280
75	AAAAATGGAA	CAGCCACCC	TTTTGCCCC	ATAGCTGTAG	TTAGAATTGA	GTACCTGTAG	8340
	CAAAACAGCT	GTAATTTGGT	GTTGTAGTGT	TAGAGGTGTT	AGCTTGCTAG	TGACTAGCTT	8400
	TGGAGAGTAA	ATGCATGGTA	TTGTACATCA	CATTTCTTAA	CTGTTTTTAA	CCTCTGAAA	8460
	GAATATATTC	TTCTTTGTAG	TCCTTCTTCC	CACCCCTTGG	CCCTCTCCT	CTCCCTGCTC	8520
	CCAGTTGTCT	TACAGTTGTA	AATATCTGAT	TTGAGGCCCA	ATAACTCTTG	CCAAGTAAAG	8580
	TCAGCAAAACA	ACAAACAAAC	CAAAATGTGG	GGAAAAGGCA	TTTCTCAACC	ATCTCTCAGC	8640
80	AGTTATTGAT	AGTTTCTTAA	GGAAACAGCAT	TGTGATCAAA	GACTCAACTT	TACGTAAAAA	8700
	TCAGTGGTAA	ATGGGGTTG	TATTGGCCAT	TGATTACATT	CAGGATTGAA	TAGTTTTTCA	8760
	AATCACATGT	AATCCAAAGA	CAGTAGGTAG	TGATGTCCT	TATCCCTGCA	GCTGTTTTAA	8820
	GATAGAGACC	TCAGAAGACT	CTGCTTGACC	GATGACCAAT	AATTATTGTA	AAAAAAGAA	8880
	AAAAATGAGA	GAAATAAACC	AGATATTTAA	GAACCTTAGC	CACCTATTTA	GAATAGTTAT	8940
85	AGCCAGAAAA	AAAAACAAGG	GCATGAGTTC	AAATGCAITTA	CTATCAGTGT	CCTAGGCAAT	9000
	ACCTAACCTA	CTCTGAATTT	GTGATTCAAA	AGCAGTATTT	CAAGAGGCAT	TCTCCTTTTT	9060
	TGTTTGTCTG	ACCCCACTTG	GACTGGTAGG	TTTGGTGAGG	CCCCATATAA	CCAGCTGGAG	9120

WO 02/086443

PCT/US02/12476

CAGACCCCTT TCATCTCTG TGCCCTGTAAC ACCCTCTCTC CCCCACCCCC TCCGCAATTC 9180  
AATGAGGGCT TTCTTGGGTC AGAGGAGCTTC AAGGTTGTCT AGAGAAGTTT GCCATGTGTG 9240  
TAAGGTGCTG TGAAGCTGTA GTGCTGAAGA TTGCGAGCAT TCAATACCAG GCAGCCAAAG 9300  
AGCTGCTCTT GCAATTATTT TGGCTCTCAA GCTCTGTCTT TCATCGCATT CTCATTTCTG 9360  
TGTACATTG CAAGATGTGT GTAATGTCAT TTTCCAAAAA TAAATTTGA TTTCAAT

Seq ID NO: 214 Protein sequence:  
Protein Accession #: NP\_000546

1 11 21 31 41 51  
MELDFGHFDE RDKTSRNMNRG SRMNGLPSP T HSAHCSFYRT RTLQALSNEK KAKIKVRFYRN 60  
GDRYFKGIVY AVSSDRFRSF DALLADLTRS LSDNINLPQG VRYIYITDGS RKIGSMDELE 120  
EGESYVCSDD NFFKKVEYTK NVNPNWSVNV KTSANMKAPQ SLASSNSAQA RENKDFVRPK 180  
LVTIIRSGVK PRKAVRVLLN KTAHSFEQV LTDITEAIKL ETGVVKKLYT LDGKQVTCLEH 240  
DFFGDDVFI ACPGPEKFRYA QDDPSLDENE CRVMKGNPSA TAGPKASPTF QKTSKSPGP 300  
MRRSKSPADS ANGTSSSLSL TPKSKQSPIS TPTSPGSLRK HEDLYLPLSL DSDSLGDSM

Seq ID NO: 215 DNA sequence  
Nucleic Acid Accession #: NM\_130467  
Coding sequence: 312..644

1 11 21 31 41 51  
GGCAGGAGGC AGAGCTCTGC AAGGAGAGGT TGTGTCTTCG TTTCTTCCGC CATCTTCGTT 60  
CTTTCACACA TCTTCGTCTT TTCTCACTGA CCGAGACTCA GCGGTAGGT CTGCAGAGTG 120  
GTCTTCCTGG TAAATTAGTT GTGAGTGAAT GTGTGGAGGA GCCAGCGGC TTAGACAGG 180  
TCTGTGSCA CAGTCCGTGG CTTTGAAGGA AAAGGGCCTC GCGGTGTGTC TCCGCTCTCC 240  
CCGAGTCTGT GATGCAGGCG CCATGGGCGG GTAATCGTGG CTGGGTCTGA ACGAGGGAGC 300  
AAGTGAGAGA TATGAGTGAG CATGTAACAA GATCCCAATC CTCAGAAAGA GGAATGACC 360  
AAGAGTCTTC CCACCCAGTT GGAGCTGTGA TTGTCCAGCA GCCCACTGAG GAAAAACGTC 420  
AAGAAGAGGA ACCACCAACT GATAATCAGG GTATTGCACC TAGTGGGAG ATCAAAAATG 480  
AAGGAGCACC TGCTGTTCAT GGGACTGATG TCGAAGCTTT TCAACAGGAA CTGGCTCTGC 540  
TTAAGATAGA GGATGCACCT GGAGATGGTC CTGATGTGAG GGAGGGGACT CTGCCCACTT 600  
TTGATCCAC TAAAGTGTCT GAAGCAGGTG AAGGGCAACT ATAGGTTTAA ACCAAGACAA 660  
ATGAAGACTG AAACCAAGAA TATGTGTTCT ATGCTGGAAT TTTGACTGCT AACATTCTCT 720  
TAATAAAGTT TTACAGTTT CTGCAAAAAA AAAAAAAAAA AAA

Seq ID NO: 216 Protein sequence:  
Protein Accession #: NP\_569734

1 11 21 31 41 51  
MSEHVTRSQS SERGNDQESS QPVGPIVQQ PTEKRQEEB PPTDNQGIAP SGEIKNEGAP 60  
AVQGTDEAF QQELALLKIE DAPGDGPDVR BGTLPFDFT KVLEAGEGQL

Seq ID NO: 217 DNA sequence  
Nucleic Acid Accession #: NM\_001476.1  
Coding sequence: 82..435

1 11 21 31 41 51  
GCCAGGGAGC TGTGAGGAGC TGCTGTGTGG TTCTGCGCT CCGGACTCTT TTTCTCTAC 60  
TCAGATTCAT CTGTGTGAAA TATGAGTTGG CGAGGAAGAT CGACCTATTA TTGGCTTAGA 120  
CCAAGGCGCT ATGTACAGCC TCCTGAAGTG ATTGGGCTTA TGCGGCCCGA GCAGTTCACT 180  
GATGAAGTGG AACCAACAAC ACCTGAAGAA GGGGAACCA CAACTCAAGC TCAGGATCCT 240  
GCAGCTGCTC AGGAGGGAGA GGATGAGGGA GCATCTGCAG GTCAAGGGCC GAAGCCTGAA 300  
GCTGATAGCC AGGAACAGGG TCACCCACAG ACTGGGTGTG AGTGTGAAGA TGTCTCTGAT 360  
GGGAGGAGG TGGACCCGCC AAATCCAGAG GAGGTGAAAA CGCCTGAAGA AGGTGAAAAG 420  
CAATCACAGT GTTAAAGAA GACACGTTGA AATGATGCAG GCTGCTCTTA TGTGGAAAT 480  
TTGTTCAATTA AATTTCTCCC AATAAAGCTT TACAGCCTTC TGCAAAA

Seq ID NO: 218 Protein sequence:  
Protein Accession #: NP\_001467.1

1 11 21 31 41 51  
MSWRGRSTYY WPRPRRYVQP FEVIGPMRPE QFSDEVERAT PEEGEFATOR QDPAAQEGE 60  
DEGASAGQGP KPEADSQEQG HPQTGCEDD GPDGQEVDFP NPEEVKTPPE GERQSQS

Seq ID NO: 219 DNA sequence  
Nucleic Acid Accession #: NM\_001476  
Coding sequence: 90-3671

1 11 21 31 41 51  
ACAGCGGAGC GCAGAGTGAG AACCACCAAC CGAGGCGCGG GGCAGCGACC CTTGCAGCGG 60  
AGACAGAGAC TGAGCGGCCG GGCACCGCCA TGCTGCGCT CTGGCTGGGC TGCTGCCTCT 120  
GCTTCTGCTT CCTCTGCCC GCAGCCCGGG CCACCTCCAG GAGGGAAGTC TGTGATTGCA 180  
ATGGGAAGTC CAGGCACTGT ATCTTTGATC GGGAACTTCA CAGACAACT GGTAAATGGAT 240  
TCCGCTGCTT CAATGCAAT CACAACATG ATGGCATTCA CTGCGAGAAG TGCAAGAAATG 300  
GCTTTTACCG GCACAGAGAA AGGGACCGCT GTTTGCCCTG CAATTGTAAC TCCAAAGGTT 360



	CTCTTAGTGC	TCGATGTGAC	AACTCTGGAC	GGTGCGAGCTG	TAAACCAGGT	GTGACAGGAG	420
	CCGATGTGCGA	CCGATGTGCTG	CCAGGCTTCC	ACATGCTCAC	GGATGCGGGG	TGCACCCCAAG	480
	ACCAGAGACT	GCTAGACTCC	AAGTGTGACT	GTGACCCAGC	TGGCATCGCA	GGGCGCTGTG	540
5	ACGCGGCGCG	CTGTGTCTGC	AAGCCAGCTG	TACTTGAGAG	ACGCTGTGAT	AGGTGTGATG	600
	CAGGTTACTA	TAATCTGGAT	GGGGGGAACC	CTGAGGGCTG	TACCCAGTGT	TTCTGTCTATG	660
	GGCATTTCAGC	CAGCTGCCCG	AGCTCTGCAG	AATACAGTGT	CCATAAGATC	ACCTCTACCT	720
	TTTATCAAGA	TGTTGATGGC	TGGAAGGCTG	TCCAACGAAA	TGGGTCTCCT	GCAAAAGCTCC	780
	AATGTCACAC	CGGCCATCAA	GATGTGTTTA	GCTCAGCCCA	ACGACTAGAC	CCTGTCTATT	840
10	TTGTGGCTCC	TGCCAAATTT	CTTGGGAATC	AACAGGTGAG	CTATGGGCAA	AGCCTGTCCCT	900
	TTGACTACCG	TGTGGACAGA	GGAGGCAGAC	ACCCATCTGC	CCATGATGTG	ATTCTGGAAG	960
	GTGCTGTGCT	ACGGATCACA	GCTCCCTTGA	TGCCACTTGG	CAAGACACTG	CCTTGTGGGC	1020
	TCACCAAGAC	TTACACATTC	AGGTTAAATG	AGCATCCAAG	CAATAATTGG	AGCCCCCAGC	1080
	TGAGTTACTT	TGAGTATCGA	AGGTTACTGC	GGAACTCTAC	AGCCCTCCGC	ATCCGAGCTA	1140
	CATATGAGAG	ATACAGTACT	GGGTACATTG	ACAAATGTAG	CCTGATTTC	GCCCCCCTG	1200
15	TCTCTGGAGC	CCGAGCACC	TGGGTGAAC	AGTGTATATG	TCCTGTGGG	TACAAGGGGC	1260
	AATTCTGCCA	GGATTGTGCT	TCTGGCTACA	AGAGAGATTC	AGCGAGACTG	GGGCCTTTTG	1320
	GCACCTGTAT	CTTTGTAAC	TGTCAAGGGG	GAGGGGCGCTG	TGATCCAGAC	ACAGAGAGATT	1380
	GTTATTGAGG	GGATGAGAAT	CCTGACATTG	AGTGTGCTGA	CTGCCCAATT	GGTTTCTACA	1440
20	ACGATCCGCA	CGACCCCGC	AGCTGCAAGC	CATGTCCCTG	TCATAACGGG	TTACAGTGCT	1500
	CAGTGATGCC	GGAGCCGGAG	GAGGTGGTGT	GCAATAACTG	CCCTCCCGGG	GTACCGGGTG	1560
	CCCGCTGTGA	GCTCTGTGCT	GATGGCTACT	TTGGGGACCC	CTTTGGTGAA	CATGGCCCG	1620
	TGAGGCTTGT	TCAGCCCTGT	CAATGCAACA	ACAATGTGGA	CCCCAGTGCC	TCTGGAATT	1680
	GTGACCGGCT	GACAGGCGAG	TGTTTGAAGT	GTATCCACAA	CACAGCCGGC	ATCTACTGCG	1740
25	ACCACTGCAA	ACGAGGCTAC	TTCCGGGACC	CAITGGCTCC	CAACCCAGCA	GACAAAGTTC	1800
	GAGCTTGCAA	CTGTAAACCC	ATGGGCTCAG	AGCCCTGTAGG	ATGTCGAAGT	GATGGCACC	1860
	GTGTTTGCAA	CCGAGGATT	GGTGGCCCCA	ACTGTGAGCA	TGGAGCATTC	AGCTGTCCAG	1920
	CTTGCTATAA	TCAAGTGAAG	ATTGAGATGG	ATCAGTTTAT	GCAGCAGCTT	CAGGAATGG	1980
	AGGCCCTGAT	TTCAAAGGCT	CAGGCTGGTG	ATGGAGTAGT	ACCTGATACA	GAGCTGGAAG	2040
30	GCAGGATGCA	CGAGGCTGAG	CAGGCCCTTC	AGGACATTCT	GAGAGATGCC	CAGATTTTCA	2100
	AAGGTGTGAG	CAGATCCCTT	GGTCTCCAGT	TGGCCAAAGT	GAGGAGCCAA	GAGAACAGCT	2160
	ACCAAGGCGG	CCTGGATGAC	CTCAAGATGA	CTGTGGAAG	AGTTCCGGCT	CTGGGAAGTC	2220
	AGTACCAGAA	CGAGTTCGCG	GATACTCACA	GGCTCATCAC	TCAGATGCAG	CTGAGCCTGG	2280
	CAGAAAGTGA	AGCTTCCTTG	GGAAACACTA	ACATTCTCTG	CTCAGACCAC	TACGTGGGCG	2340
35	CAAAATGGCTT	TAAAGTCTG	GCTCAGGAGG	CCACAAGATT	AGCAGAAAGC	CACGTTGAGT	2400
	CAGCCAGTAA	CTAGGAGCAA	CTGACAAAGG	AAACTGAGGA	CTATTCCAAA	CAAGCCCTCT	2460
	CACTGGTGCG	CAGGCCCCTG	CATGAAGGAG	TGGAAAGCGG	AAGCGTAGC	CCGAGCGGTG	2520
	CTGTGGTCCA	AGGGCTGTGT	GAAAAATTGG	AGAAAAACCA	GTCCCTGGCC	CAGCAGTTGA	2580
	CAAGGGAGGC	CACTCAAGCG	GAAATGAAG	CAGATAGGTC	TTATCAGCAC	AGTCTCCGCC	2640
40	TCTTGATTTC	AGTGTCTCGG	CTCAGGGAG	TCAGTGATCA	GTCTTTTCAG	GTGGAAGAG	2700
	CAAGAGGAGT	CAACCAAAA	GCGGATTCAC	TCTCAACGCT	GGTAACGAGG	CATATGGATG	2760
	AGTTCAAGCG	TACACAAAAG	AATCTGGGAA	ACTGGAAAGA	AGAAGCACAG	CAGCTCTTAC	2820
	AGAATGGAAA	AAGTGGGAGA	GAGAAATCAG	ATCAGCTGCT	TTCCCGTGCC	AATCTTGCTA	2880
	AAAGCAGAGC	ACAAGAAACA	CTGAGTATGG	GCAATGCCAC	TTTTTATGAA	GTGAGAGCA	2940
45	TCCTTAAAAA	CCTCAGAGAG	TTTGACCTGC	AGGTGGACAA	CAGAAAAGCA	GAAGCTGAAG	3000
	AGGCCATGAA	GAGACTCTCC	TACATCAGCC	AGAAGGTTTC	AGATGCCAGT	GACAAGACCC	3060
	AGCAAGCAGA	AAGAGCCCTG	GGGAGCGCTG	CTGCTGATGC	ACAGAGGGCA	AAGAATGGGG	3120
	CCGGGGAGGC	CCTGGAAATC	TCCAGTGAGA	TTGAACAGGA	GATTGGGAGT	CTGAACTTGG	3180
	AAGCCAAATG	GACAGCAGAT	GGAGCCTTGG	CCATGGAAAA	GGGACTGGCC	TCTCTGAAGA	3240
50	GTGAGATGAG	GGAGTGGGAA	GGAGAGCTGG	AAAGGAAGGA	GCTGGAGTTT	GACACGAATA	3300
	TGGATGCACT	ACAGATGTGT	ATTACAGAAG	CCCAGAAGGT	TGATACCAGA	GCCAAAGAAC	3360
	CTGGGGTTAC	AATCCAAGAG	ACACTCAACA	CATTAGACGG	CCTCCTGCAT	CTGATGGACC	3420
	AGCCTCTCAG	TGTAGATGAA	GAGGGGCTGG	TCTTACTGGA	GCAGAACTT	TCCCGAGCCA	3480
	AGACCCAGAT	CAACAGCCAA	CTGCGGCCCA	TGATGTGAGA	GCTGGAAAGG	AGGGCAGGTC	3540
55	AGCAGAGGGG	CCACCTCCAT	TTGCTGGAGA	CAAGCATAGA	TGGGATTCTG	GCTGATGTGA	3600
	AGAACTTGGA	GAACTATTAG	GACAACTTGC	CCCCAGGCTG	CTACAATACC	CAGGCTCTTG	3660
	AGCAACAGTG	AAGCTGCCAT	AAATATTCT	CAACTGAGGT	TCTTGGGATA	CAGATCTCAG	3720
	GGCTCGGAG	CCATGTCATG	TGAGTGGGTG	GGATGGGGAC	ATTGGAACAT	GTTTAATGGG	3780
	TATGCTCAGG	TCAACTGACC	TGACCCCAT	CCTGATCCCA	TGGCCAGGTG	GTGTCTTAT	3840
60	TGCACCATAC	TCTTGTCTTC	CTGATGCTGG	GCAATGAGGC	AGATAGCACT	GGGTGTGAGA	3900
	ATGATCAAGG	ATCTGGAGCC	CAAAGAATAG	ACTGGATGGA	AAGACAAACT	GCACAGGCAG	3960
	ATGTTTGCTT	CATAATAGTC	GTAAGTGGAG	TCTTGGAAAT	TGGACAAGTG	CTGTTGGGAT	4020
	ATAGTCAACT	TATCTTTTGA	GTAATGTGAC	TAAAGGAAAA	AACITTGACT	TTGCCAGGGC	4080
	ATGAATTTCT	TCTTAATGTC	AGAACAGAGT	GCAACCCAGT	CACACTGTGG	CCAGTAAAT	4140
65	ACTATTGCCCT	CATATTGTCC	TCTGCAAGCT	TCTTGCTGAT	CAGAGTTCTT	CCTACTTACA	4200
	ACCCAGGGTG	TGAACATGTT	CTCCATTTTC	AAGCTGGAAG	AAGTGAGCAG	TGTTGGAGTG	4260
	AGGACCTGTA	AGGCGAGGCC	ATTGAGAGCT	ATGGTGTCTG	CTGGTGCCCTG	CCACCTTCAA	4320
	GTCTTGAGCC	TGGGCATGAC	ATCCTTTCTT	TAAATGATGC	CATGGCAACT	TAGAGATTGC	4380
	ATTTTATTATA	AAGCATTTCC	TACCAGCAAA	GCAAAATGTTG	GGAAAGTATT	TACTTTTTCG	4440
70	GTTTCAAAGT	GATAGAAAAA	TGTGGCTTGG	GCAATTGAAA	AGGTAAAATT	CTCTAGATTT	4500
	ATTAGTCCCTA	ATTCAATCCT	ACTTTTCGAA	CACCAAAAAAT	GATGCGCATC	AATGTATTTT	4560
	ATCTTATTTT	CTCAATCTCC	TCTCTCTTTC	CTCCACCCAT	AATAAGAGAA	TGTTCTTACT	4620
	CACACTTCAG	CTGGGTCAAC	TCCATCCCTC	CATTTCATCT	TCCATCCATC	TTTCCATCCA	4680
	TTACCTTCAT	CCATCCTTCC	AACATATATT	TATTGAGTAC	CTACTGTGTG	CCAGGGGCTG	4740
75	GTGGGACAGT	GGTGACATAG	TCTCTGCCCT	CATAGAGTTG	ATTGTCTAGT	GAGGAAGACA	4800
	AGCATTTTFA	AAAAATAAAT	TTAAACTTAC	AAACTTTGTT	TGTCACAAGT	GGTGTTTATT	4860
	GCAATAACCG	CTTGGTTTGC	AACCTCTTGG	CTCAACAGAA	CATATGTTGC	AAGACCTTCC	4920
	CATGGGGGCA	CTTGAGTTT	GGCAAGGCTG	ACAGAGCTCT	GGGTGTGCA	CATTTCCTTG	4980
	CATTCCAGCT	GTCACCTCTG	GCCTTCTAC	AACTGATTGC	AACAGACTGT	TGAGTTATGA	5040
80	TAAACACAGT	GGGAATGCT	GGAGGAACCA	GAGGCACCTC	CACCTTGGCT	GGGAAGACTA	5100
	TGCTGTCTGC	TTGCTCTGCT	ATTTCCCTGG	ATTTTCCTGA	AAGTGTTTT	AAATAAAGAA	5160
	CAATTGTTAG	ATGCC					

Seq ID NO: 220 Protein sequence:  
Protein Accession #: NP\_005553

1 11 21 31 41 51  
| | | | |

WO 02/086443

PCT/US02/12476

MPALWLGCCCL CFSLLLPAAAR ATSRREVCDG NGKSRQCIFD RELHROQTNG FRCNLNCNDNT 60  
DGIHCEKCKN GPFYRHRERRR CLPCNCKNSG SLSARCDNSG RCSCKPGVTG ARCDRCCLPGF 120  
HMLTDAGCTQ DQRLDSKCD CDPAGIAGPC DAGRCVCKPA VTGERCDRCR SGYYNLDGDN 180  
PEGCTQCFY GHASCRSSA EYSVHKITST FHQDVGWKA VQRNGSPAKL QWSQRHQDV 240  
SSAQRDLPVY FVAPAKFLGN QQVSYGQSLF FDYRVDGRGR HPSAHDVILE GAGLRITAPL 300  
MPLGKTLPCG LTKTYTFRLN EHPNNWSPQ LSYFEYRLL RNLTLALIRA TYGEYSTGYI 360  
DNVTLISARP VSGAPAPWVE QCICPVGYKG QFCQDCASGY KRDSARLPGF GTCIPCNOQG 420  
GGACDPDTGD CYSGDENPDI ECADCPGFY NDHDPDRCK PCPCHNGPSC SVMPEZEEVV 480  
CNCPPPGVTG ARCELCAAGY FGDPPFGEHGF VRPCQPCQCN MNVDPSASGN CDRLTGRCLK 540  
CIHNTAGIYC DQCKAGYFGD PLAPNPADKC RACNCPMGS EPVGCSDGT CVCKPGFGGP 600  
NCEHGAFCSC ACYNQVKIQM DQPMQQLQRM EALISKAQGG DGVVPDTELE GRMQQAEQAL 660  
QDILRDAQIS EGASRSLGLQ LAKVRSQENS YQSRLLDLKM TVERVRLGSG QYQNRVRDTH 720  
RLITQMQLSL AESEASLGNT NIPASDHYVG PNGFKSLAQE ATRLAESHVE SASNMEQLTR 780  
ETEDYSKQAL SLVRKALHEG VSGSGSPFDG AVVQGLVEKL EKTSLAQQL TREATQABIE 840  
ADRSYQHSIR LLDVSRLQG VSDQSPQVEE AKRIKQADS LSTLVTRHMD EFKRTQKNLG 900  
NWKEEAQQLL QNGKSGREKS DQLLSRANLA KSRQAQALSM GNATPYEVES ILKNLREFDL 960  
QVDNRKAEAE EAMKRLSYIS QKVSADSKT QQAERLQSA AADAQRAKNG AGEALBISSE 1020  
IEQSIGSLNL EANVTADLGL AMEKGLASLK SEMREVEGEL ERKELEFDTN MDAVOMVITE 1080  
AQKVDTRAKN AGVTIQDTLN TLDGLLHMD QPLSVDBEGL VLEQKLSRA KTQINSQLRP 1140  
MMSLEERAR QQRGHLHLE TSIDGILADV KNLENIRDNL PPGCYNTQAL EQQ

Seq ID NO: 221 DNA sequence  
Nucleic Acid Accession #: NM\_016529  
Coding sequence: 13-1854

1 11 21 31 41 51  
GTCAAGAAAA GAATGCTCTGT AATTGTTCTGA ACTCCTTCAG GACGACTTCG GCTTTACTGT 60  
AAAGGGGCTG ATAATGTGAT TTTTGAGAGA CTTTCAAAAG ACTCAAATA TATGGAGGAA 120  
ACATTATGCC ATCTGGAATA CTTTCCACG GAAGGCTTGC GGACTCTCTG TGTGGCTTAT 180  
GCTGATCTCT CTGAGAAATGA GTATGAGGAG TGGCTGAAAG TGTATCAGGA AGCCAGCACC 240  
ATATTGAAGG ACAGAGCTCA ACGGTTGAA GAGTGTACG AGATCATTGA GAAGAATTG 300  
CTGCTACTTG CAGCCACAGC CATAGAAGAT CGCCTTCAAG CAGGAGTTC AGAAACCATC 360  
GCAACACTGT TGAAGGCAGA AATTAAATA TGGGTGTGA CAGGAGACAA ACAAGAACT 420  
GCGATTAATA TAGGCTATTC CTGCCGATTG GTATCGCAGA ATATGCGCCT TATCCTATTG 480  
AAGGAGGACT CTTTGGATGC CACAAGGGA GCCATTACT AGCACTGCAC TGACCTTGGG 540  
AATTGCTCGG GCAAGGAAAA TGACGTGGCC CTCATCATCG ATGCCACAC CCTGAAGTAC 600  
GCGCTCTCCT TCGAAGTCCG GAGGAGTTTC CTGGATTGG CACTCTCGTG CAAGACGCTC 660  
ATATGCTGCA GAGTGTCTCC TCTGAGAGG TCTGAGATAG TGGATGTGGT GAAGAGCGG 720  
GTGAAGGCCA TCACCCCTGC CATCGGAGC GGCGCCACG ATGTCCGGAT GATCCAGACA 780  
GCCACGCTG GTGTGGGAAT CAGTGGGAAT GAAGGCATGC AGGCCACCA CACTCGGAT 840  
TACCCATCG CACAGTTTTT CTACTTAGAG AAGCTTCTGT TGGTTCATGG AGCCTGGAGC 900  
TACAACGGG TGACCAAGTG CATCTGTATC TGCTTCTATA AGAACGTGGT CCTGTATATT 960  
ATTGAGCTTT GGTTCGCTT TGTTAATGGA TTTTCTGGG AGATTTTATT TGAACGTTGG 1020  
TGCTACGGCC TGTAACAATG GATTTTCAAC GCTTTGCCG CCTCACTCT GGAATCTTT 1080  
GAGAGGCTCT GCATCAGGA GAGCATGCTC AGGTTTCCC AGCTCTACAA AATCACCCAG 1140  
AATGGGAGG GCTTCAACAC AAAGGTTTTC TGGGTCATC GCATCAACGC CTGTGTCAC 1200  
TCCCTCATCC TCTTCTGTT TCCCATGAAA GCTCTGGAGC ATGATCTGT GTTTGACAGT 1260  
GGTCATGCTA CCGACTATT ATTTGTGGA AATATTGTT ACACATATGT TGTGTATTCT 1320  
GTTTGTCTGA AAGCTGTTT GGAGAACACA GCTTGAACA AATTCAGTCA CTCGGCTGTC 1380  
TGGGGAAGCA TGCTGACCTG GCTGGTGT TTGGCATCT ACTCGACCAT CTGGCCACC 1440  
ATTCCCATG TCCCATGAT GAGAGGACAG GCAACTATG TCTGAGCTC CGCACACTTC 1500  
TGGTTGGGAT TATTCTGGT TCCTACTGCC TGTTTGATT AAGATGTGGC ATGGAGAGCA 1560  
GCCAAGCACA CCTGCAAAA GACATTGCTG GAGGAGGTGC AGGAGCTGGA AACCAAGTCT 1620  
CGAGTCTCGG GAAAGCGGT GCTGCGGAT AGCAATGGA AGAGGCTGAA CGAGCGCGAC 1680  
CGCTGATCA AGAGGCTGGG CCGGAAGAG CCCCCAGC TGTTCGGGG CAGCTCCCTG 1740  
CAGCAGGGG TCCCGCATG GTATGCTTT TCTCAAGAAG AACACGGAGC TGTAGTTCAG 1800  
GAAGAAGTCA TCGGTGCTTA TGACACCACC AAAAAGAAAT CCAGGAAGAA ATAAGACATG 1860  
AATTTCTCTG ACTGATCTTA GGAAGAGAT TCAGTTTGT GCACCCAGTG TTAACACATC 1920  
TTTGTGAGAG AAGACTGGCG TCCAAGGCCA AAACACCAGG AAACACATTT CTGTGGCCTT 1980  
AGTTAAGCAG TTTGTAGTT ACATATCCC TCGCAACCT GAGTGCAGA CCACAGGGGA 2040  
AGCTATCTTT GCCCTCCCAA CTCGTCTGCA GTGCTTAGCC TAACCTTTGT TTAGTCTGTT 2100  
ATGAAGCATT CAACTGTGCT CTGTGAGGTC TCAAATTAAT AACATTATGT TTCACCAATA 2160  
AGAAAAAAA AAAAAAA

Seq ID NO: 222 Protein sequence:  
Protein Accession #: NP\_057613

1 11 21 31 41 51  
MSVIVRTPSG RLRLYCKGAD NVIFERLSKD SKYMEETLCH LEYFATEGLR TLCVAYADLS 60  
ENEYEELKV YQRASTILKD RAORLEECYE IIEKNLLLLG ATAIEDRLQA GVPETIATLL 120  
KAEIKIWLVT GDKQETAINI GYSCRLVSN MALILLKEDS LDATRAAITQ HCTDLGNLLG 180  
KENDVALIID GHTLKYALSF EVRRSFLDLA LSKKAVICCR VSFLQKSEIV DVVKRVRKAI 240  
TLAIGDGAID VGMQTAHVQ VGISGNEGMO ATNNSDYAIA QFSYLEKLLL VHGAWSYNRV 300  
TKCILYCFYK NVVLYIIELW FAPVNGFSQG ILFERWCIGL YNVIFTALPP FTLGIFERS 360  
TQESMLRFPQ LYKITQNGEG FNTKVFNGHC INALVHSLIL FWFPMKALEH DTVFDSGHAT 420  
DYLFVGNIVY TYVVVTVCLK AGLETTAWTK FSHLAVWGSN LTLWLVFFGIY STIWFTPIA 480  
PDMRGQATMV LSSAHFWLGL FLVPTACILIE DVANRAAKHT CKKTLLEEVO ELETKSRVLG 540  
KAVLRDSNGK RLNERDRLIK RLGRKTPPTL FRGSSLQGGV PHGYAFSQEE HGAVSQEVI 600  
RAYDTTKKRS RKK

Seq ID NO: 223 DNA sequence  
Nucleic Acid Accession #: BC017001  
Coding sequence: 1-394

1 11 21 31 41 51

5 AACGCTGGGC AGGCGCCGCG CGGGTCGGGG GGGCCCCGAG GGGCCCCGGC CGAGCGGCGG 60  
 CGCGCAGGGC GGCAGCATCC ACTCGGGCCG CATCGCCGCG GTGCACAAAG TGCCGCTGAG 120  
 CGTGTCTATC CGGCGCGTGC CGTCCGTGTT GGACCCCGCC AAGGTGCAGA GCCTCGTGA 180  
 CACGATCCGG GAGGACCCAG ACAGCGTGCC CCCCATCGAT GTCCCTCTGA TCAAGGGGCG 240  
 CCAGGGAGGT GACTACTTCT ACTCCTTTGG GGGCTGCCAC CGCTACGCGG CCTACCGACA 300  
 ACTGCGCGGA GAGACCATCC CGCCCAAGCT TGTCCAGTCC ACTCTCTCAG ACCTAAGGGT 360  
 GTACCTGGGA GCATCCACAC CAGACTTGCA GTAGCAGCCT CCTTGGCACC TGCTGCCACC 420  
 10 TTCAAGAGCC CAGAAGACAC ACCTGGCCTC CAGCAGGCTG GGCCATGCAG AAGGGATAGC 480  
 AGGGGTGCAT TCTCTTTGCA CCTGGCGAGA GGGTCTGACT CTGGGCACCC CTCTCACCGG 540  
 CTACAAGGCC TTGGACTCAC TGTACAGTGT GGGAGCCCCA GTTCCCACCT CTGTGACAAT 600  
 AGGATCATGG CCTTACCCTT GAAGCATTAC CGAAGAGGAG AACAGAGATG GGCTTGAAGA 660  
 GCCACGTGCT GCGCGCTCCA AATTCCCAAG GACAAGGATC CCTCTGCATT TTGTCTATG 720  
 15 TAACTCTTA TAGGACTAC ATTCACTGTC AAGGAAAGGA AAACCTTGAT TGCAGTGGTT 780  
 TAAACAAACA GAAGATGTTT TTCCACATA GCATGGATTG TGGAGATGGG TGGCTAATGG 840  
 TATTGGTTCA ACAACTCCAC GGAGGTAGGG GTCACTCTT GGATCCTTTT GCCTTAATCT 900  
 CAGTGTCTGT TACTTCATGG TCCCAAGATG GCTGCTGTAT CCCCAGAAT CATGTCTGCG 960  
 20 TTCAAGGAAG GAGGGTGGTA GGAAGAGGAA GGGCCAAACT AGCTGGACCC GTCCCTTCT 1020  
 ATCAGAAAGT AAACCTCTGT CAGAAGTCTG TTCTCTGCTC TCTCCCTCTG CATATCTTCA 1080  
 CTTAGATGCC CTTGGCCCGA GCCAGCTACC ATTGCACCTC TAGCTGCAAA CAAAGCTAAG 1140  
 ACAGCAGGGA ACAGAAATGT CATGGCTGAA TAGACCAATC GTTTCATC TACTGAGACT 1200  
 GGCACACTGC CTCCTGCAAT AAACTGGGA TCCCATTACC AAGAGAGAAA TGCAGAATTG 1260  
 TGTACAGTT AGCTTTTGCT GTGTAACAAA CCATCCCCAA ACTTGGCAGC TAGAAACAAA 1320  
 25 CCCTTATTT TCCCAAAATC CTATGGGTTG GCAATTTGGG CTGGGCTCAA CAGGGCAGTT 1380  
 CTGCTGTCTA CACCTGGGAT CCCTCATGGA GCTAAGGTCA GCTGTACCT CAGCTGGGCC 1440  
 TGGATGGTCT AGSATAGCCT TACTCACTTG CCTGGCAGGT GACAGGCTGT TGGCTGAAT 1500  
 TGCTTGGTTC TGCTCCATGT GGCTCTCCA GCAGCTAGC TCAGGCTTAT TCACATGATG 1560  
 GCTTCAGGAT TCCAAAGAGA GTGAGAGTAG AAGCTGAAAG ACTTCTTGAG TTCTTGGCCT 1620  
 30 GGAATCGGGA GTAGACAGAT GTCACCTCTG CTAAGTTCTT TTGGTCAGAG CAAATCACAA 1680  
 GGCTTACCAC AGATTCAAGG GATGAGAAAC AGACTACATG TCTTGATGAG GGAACCCACA 1740  
 AAGAGCTTGT GGCCATTTT CACCTATCAC AAATAATTTT GGATGGGTAT TTATTGGAT 1800  
 AAAGGTATTT CCCTCTTCCC CTTTCTCTC TGTCTCATGG GGCTCACTC TGCCAAGTTG 1860  
 GAAGGCATA AGACATTGTC CTGGCCCTCA GGGTCTAGGG GAAGAGGTGT TGGGCGAGGA 1920  
 35 AGTGAGTCTC TCCATGGGCT GGACCCACTG TAGTAGGAGT GCCTCCTGT CTGCACTGCT 1980  
 GGTATGGGGT TAGGCCAGGT AGGACATTCC AGAGGGGCTT CTGAAACCA AGATCCCTG 2040  
 GGGAAAGGGA ACAGATGAAG GCAGGCTTGG TTCTCACTGC CCTCTAAGGG AACTTGTCA 2100  
 CTGGGCATT TTAAGCCTCA GTTCTCCAG TTCAATAATA AGGACAAGAG CTTTCCCAT 2160  
 GCATTCTCTT TCCCGGGGAA AGTTGACTGA GGTGACCAAT AATAGAATTG AAAAGGGAGA 2220  
 40 GTGTCTCAG TGCAATGTGG CATCTCGAT TGGGTCTTGG AACAAAACA GGACATTAGT 2280  
 GGGAAATTTG GAAATCTGAA AAAAGTCTGA ATTTAGTTA ATATACCAAT TTCAGTCTCT 2340  
 TGGTTTTCAG AGATGTACCA TGGTGATGTA AGATGTTGAC CTTGGGGTAG GCTGGGTGAA 2400  
 GGGTATACAG GAACCTCTTG TACTATCTCT GCAACTTCTC TGTAAATCTA GTATCATTCC 2460  
 AAAATAAAG TTATTATAT TTAATAAATA AAAAAAAAAA AA

Seq ID NO: 224 Protein sequence:  
 Protein Accession #: AAH17001.1

50 1 11 21 31 41 51  
 TLGRAGAGRG APEGPSPSGG AQGGSIHSGR IAAVHNVPIS VLIRPLPSVL DPAKVQSLVD 60  
 TIREDDPSVP PIDVLWIKGA QGGDYFYBFG GCHRYAAYQQ LQRETIPAKL VQSTLSDLRV 120  
 YLGASTPDLQ

Seq ID NO: 225 DNA sequence  
 Nucleic Acid Accession #: NM\_021048  
 Coding sequence: 1..1110

60 1 11 21 31 41 51  
 ATGCCTCGAG CTCCAAAGCG TCAGCGCTGC ATGCCTGAAG AAGATCTTCA ATCCCAAAGT 60  
 GAGACACAGG GCTCGAGGG TGCAACGGCT CCCCTGGCTG TGGAGGAGGA TGCTTCATCA 120  
 65 TCCACTTCCA CCAGCTCCTC TTCTCCATCC TCTTTCCCTC CCTCCTCCTC TTCTCCTCC 180  
 TCCTCTCTGT ATCCTCTAAT ACCAAGCACC CCAGAGGAGG TTTCTGCTGA TGATGAGACA 240  
 CCAAATCCTC CCAGAGTGGC TCAGATAGCC TGCTCCTCCC CCTCGGTGCT TGCTTCCCTT 300  
 CCATTAGATC AATCTGATGA GGGCTCCAGC AGCCAAAAGG AGGAGAGTCC AAGCACCTTA 360  
 CAGGTCTCTC CAGACAGTGA GTCTTTACCC AGAAGTGAGA TAGATGAAAA GGTGACTGAT 420  
 70 TTGGTGCAGT TTCTGCTCTT CAAGTATCAA ATGAAGGAGC CGATCACAAA GGCAGAAATA 480  
 CTGGAGAGTG TCATAAAAA TTATGAAGAC CACTTCCCTT TGTGTTTAG TGAAGCCTCC 540  
 GAGTGCATGC TGCTGGTCTT TGGCATTGAT GTAAAGGAAG TGGATCCCAC TGGCCACTCC 600  
 TTTGTCTTGG TCACCTCCCT GGGCCTCACC TATGATGGGA TGCTGAGTGA TGTCCAGAGC 660  
 ATGCCAAGA CTGGCATTCT CATACTTATC CTAAGCATAA TCTTCATAGA GGGCTACTGC 720  
 75 ACCCTGAGG AGGTCACTCT GGAAGCACTG AATATGATGG GGCTGTATGA TGGGATGGAG 780  
 CACCTCATTT ATGGGGAGCC CAGGAAGCTG CTCACCAAG ATTGGGTGCA GGAAACTTAC 840  
 CTGAGTACC GGCAGGTGCC TGGCAGTGAT CCTGCACGGT ATGAGTTTCT GTGGGTCCA 900  
 AGGGCTCATG CTGAAATTAG GAAGATGAGT CTCTGAAAT TTTTGGCCAA GGTAAATGGG 960  
 AGTGATCCAA GATCCTTCCC ACTGTGGTAT GAGGAGGCTT TGAAGATGA GGAAGAGAGA 1020  
 80 GCCCAGGACA GAATTGCCAC CACAGATGAT ACTACTGCCA TGGCCAGTGC AAGTTCTAGC 1080  
 GCTACAGGTA GCTTCTCTTA CCCTGAATAA

Seq ID NO: 226 Protein sequence:  
 Protein Accession #: NP\_066386

85 1 11 21 31 41 51  
 MPRAKPRQRC MPREDLQSQS ETQGLEGAQA FLAVBEDASS STSTSSSPFS SFPSSSSSSS 60

WO 02/086443

PCT/US02/12476

SSCYPLIPST PEEVSADDET FNPPQSAQIA CSSPSVVASL PLDQSDGESS SQKEESPSTL 120  
QVLPSDESLEP RSEIDERYTD LVQFLLFKYQ MKEPITKAEI LESVIKNYED HFPLLPSEAS 180  
ECMLLVFGID VKEVDPTGHS FVLVTSGLT YDGMLESDVQS MPKTGILILI LSIIFIEGYC 240  
TPREVIWEAL NMMLGYDGMH HLIYGEPRKL LTQDMVQENY LEYRQVPGSD PARYEFLWGP 300  
RAHAEIKRMS LLKFLAKVNG SDPRSFPFLMY EEALKDEBER AQDRIATTDD TTAMASASSS 360  
ATGSFSYPE

Seq ID NO: 227 DNA sequence  
Nucleic Acid Accession #: NM\_005025.1  
Coding sequence: 82-1314

1 11 21 31 41 51  
GCGGAGCACA GTCGCGCGAG CACAAGCTCC AGCATCCCGT CAGGGGTGTC AGGTGTGTGG 60  
GAGGCTTGAA ACTGTTACAA TATGGCTTTC CTGGACTCT TCTCTTGCT GGTCTGCAA 120  
AGTATGGCTA CAGGGGCCAC TTTCCCTGAG GAAGCCATTG CTGACTTGTC AGTGAATATG 180  
TATAATCGTC TTAGAGCCAC TGGTGAAGAT GAAATATTC TCTTCTCTCC ATTGAGTATT 240  
GCTCTTGCAA TGGGAATGAT GGAAGTTGGG GCCCAAGGAT CTACCCAGAA AGAAATCCGC 300  
CACTCAATGG GATATGACAG CCTAAAAAAT GTGAAGAAT TTTCTTCTT GAAGGAGTTT 360  
TCAAAATCGT TAACTGCTAA AGAGAGCCAA TATGTGATGA AAATGCCAA TTCCTGTGTT 420  
GTGCAAAATG GATTTCATGT CAATGAGGAG TTTTGTCAA TGTGAAAAA ATATTTTAAT 480  
GCAGCAGTAA ATCATGTGGA CTTCAGTCAA AATGTAGCCG TGGCCAACTA CATCAATAAG 540  
TGGGTGGAGA ATAACACAAA CAATCTGGTG AAGATTGAG TATCCCCAAG GGATTTTGAT 600  
GCTGCCACTT ATCTGGCCCT CATTATGCT GTCTATTTC AAGGGAACTG GAAGTCGAG 660  
TTTAGCCCTG AAAATACTAG AACCTTTCT TCACTAAG ATGATGAAAG TGAAGTCCAA 720  
ATTCCAATGA TGTATCAGCA AGGAGAATT TATTATGGGG AATTAGTGA TGGCTCCAAT 780  
GAAGCTGGTG GTATCTACCA AGTCCTAGAA ATACCATATG AAGGAGATGA AATAAGCATG 840  
ATGCTGGTGC TGTCCAGACA GGAAGTTCCCT CTGTCTACTC TGGAGCCATT AGTCAAAGCA 900  
CAGCTGGTGT AAGAATGGGC AAACCTGTGT AAGAAGCAAA AAGTAGAAGT ATACCTGCC 960  
AGGTTCACAG TGGACAGGAA AATTGATTTA AAGATGTTT TGAAGGCTCT TGAATAACT 1020  
GAAATTTTCA TCAAGATGCG AATTGTGACA GGCCTCTCTG ATAATAAGGA GATTTTCTT 1080  
TCCAAGCAA TTCACAAGTC CTTCCTAGAG GTTAATGAAG AAGGCTCAGA AGCTGCTGCT 1140  
GTCTCAGGAA TGATTGCAAT TAGTAGGATG GCTGTGCTGT ATCCTCAAGT TATTGTGAC 1200  
CATCCATTTT TCTTCTTAT CAGAAACAGG AGAAGTGGTA CAATTCTATT CATGGGACGA 1260  
GTATCATCAT CTGAAACAAAT GAACACAAGT GGACATGATT TCGAAGAACT TTAAGTTACT 1320  
TTATTTGAAT AACAGGAAA ACAGTAACTA AGCACATTAT GTTTGCAACT GGTATATATT 1380  
TAGGATTGTT GTTTTACAGT ATATCTTAAG ATAATATTTA AAATAGTTCC AGATAAAAC 1440  
AATATATGTA AATTATAAGT AACTGTCAA GGAATGTTAT CAGTATTAAG CTAATGCTCC 1500  
TGTTATGTCA TTGTGTTTGT GTGCTGTTGT TTAATAATAA AGTACCTATT GAACATGTG

Seq ID NO: 228 Protein sequence:  
Protein Accession #: NP\_005016.1

1 11 21 31 41 51  
MAFLGLFSL VLSMATGAT PEEBAIDLS VMYNRLRAT GEDENILFSP LSIALAMGMM 60  
ELGAQSGTQK EIRHSMGYDS LKNGEEFSL KEFSNMVTAK ESQYVMKIAN SLFVQNGPHV 120  
NEEFLOMKMK YFNAAVNHVD FSNVAVANY INKVENNTN NLVKDLVSPR DFDAATYLAL 180  
INAVYFKGNW KSQFREPNTR TFSFTKDDDES EVQIPMVQQ GEFFYGEFSD GSNEAGGIYQ 240  
VLEIPYEGDE ISMMLVLSRQ EVPLATLEPL VKAQLVEENA NSVKKQKVEV YLPRFTVBQE 300  
IDLKDLVKAL GITEIFIKDA NLTLGLSDNKE IFLSKAIHKS FLEVNEEGSE AAAVSGMIAI 360  
SRMAVLYPQV IVDPFFFLI RNRRGTILF MGRVMHPETM NTSGHDFEEL

Seq ID NO: 229 DNA sequence  
Nucleic Acid Accession #: NM\_003695  
Coding sequence: 12-398

1 11 21 31 41 51  
CGACATCAGA GATGAGGACA GCATTGCTGC TCCTTGACGC CTGGGCTGTG GCTACAGGGC 60  
CAGCCCTTAC CCGCGCTGC CACGTGTGCA CCACTGCCAG CAATTGCAAG CATTCTGTGG 120  
TCTGCCCGGC CAGCTCTGCG TTCTGCAAGA CCACGAACAC AGTGGAGCCT CTGAGGGGGA 180  
ATCTGTGTGA GAAGGACTGT GCGGACTCGT GCACACCCAG CTACACCCCTG CAAGGCCAGG 240  
TCAGCAGCGG CACGAGCTCC ACCGAGTGTG GCCAGGAGGA CCTGTGCAAT GAGAAGCTGC 300  
ACAACGCTGC ACCCACCGC ACCGCCCTCG CCCACAGTGC CCTCAGCCTG GGGCTGGGCC 360  
TGAGCCTCCT GGCCTCATC TTAGCCCCCA GCCTGTGACC TTCCCCCAG GGAAGGCCCC 420  
TCATGCCCTT CTTTCCCTTT CTCTGGGGAT TCCACACCTC TCTTCCCAG CCGGCAACGG 480  
GGGTGCCAGG AGCCCCAGG TGAGGGCTTC CCCGAAAGTC TGGGACAGG TCCAGGTGGG 540  
CATGGAATGC TGATGACTTG GAGCAGGCC CACAGACCCC ACAGAGGATG AAGCCACCCC 600  
ACAGAGGATG CAGCCCCAG CTGCATGGAA GGTGGAGGAC AGAAGCCCTG TGGATCCCCG 660  
GATTTACAC TCCTTCTGTT TTGTTGCCGT TTATTTTGTG CTCAATCTC TACATGGAGA 720  
TAAATGATTT AAAC

Seq ID NO: 230 Protein sequence:  
Protein Accession #: NP\_003686

1 11 21 31 41 51  
MRTALLLAA LAVATGPALT LRCHVCTSSS NCKHSVVCPA SSRFCKTINT VEPLRGNLVK 60  
KDCAESCTPS YTLQGVSSG TSSTQCCQED LCNEKLHNA PRTALAHSA LSLGLALSL 120  
AVILAPSL

Seq ID NO: 231 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 126-752

1 11 21 31 41 51  
 1 CCGGGCCAGGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGGGTT GTCTGGAGC 60  
 5 AGGGGGCCAG GAATTCGTAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCTCTCAG 120  
 AGAAGATGAA GGATATCGAC ATAGGAAAAG AGTATATCAT CCCAGTCCCT GGTATAGAA 180  
 GTGTGAGGGA GAGAACACAGC ACTTCTGGGA CGCACAGAGA CCGTGAAGAT TCCAAGTTCA 240  
 GGAGAACTCG ACCGTGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA GCGAGGGGCC 300  
 TCTCTCTTGA TGCCCTCCATG CATTCTCAGC TCAGAATCCT GGATGAGGAG CATCCCAAGG 360  
 10 GAAAGTACCA TCATGGCTTG AGTGTCTGA AGCCCATCCG GACTACTTCC AAACACCAGC 420  
 ACCCAGTGGA CAATGTCTGG CTTTTTCTCT GTATGACTTT TCGTGGCTT TCTTCTCTGG 480  
 CCGGTGTGGC CCACAAGAAG GGGGAGCTCT CAATGGGAAG CCGTGGTCT CTGTCCAAGC 540  
 ACGAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGGCAAGAA GAGCTGAATG 600  
 AAGTTGGGCC AGACGCTGCT TCCCTGGAA GGGTTGTGTG GATCTTCTGC CGCACCAGGC 660  
 15 TCATCCTGTC CATCGTGTGC CTGATGATCA CGCAGCTGGC TGGCTTCTAGT GGACCAAAAT 720  
 TTCAAGTAGG CTGTATCTCT CGGTGAGAAT GAGAGAGTCA AGCTGGGCAG AATCTCTCGC 780  
 CAAGAGTTCA GCCTTCTCTT GGAGACTGCT CCATCAGTGC CGAGGTGTGT GGGAACAGGC 840  
 TTCACTGCAC CGCATCTTA CTGAGTTGCT TCACGTGAGG AAAAGGGGGC TTGGCCCTCG 900  
 TGACTCAGTT CCACATTTTG GATTGCATAC TGGAAAAGAA GCCAATCTTC TTGCTAGTAA 960  
 20 ACCAGCAACC CGGCTGTATA CAGTGGTGAC CCAAGCAATG GATATAAACC TAAAAATCTG 1020  
 AGGGAGGGGA GAGGTGGAAT ACAGTAGTTC TTGGAATCTG AAGTCTCCTA TTGATCAGO 1080  
 TTATTTCTCG GGACTTGGCA AAAATCTGAT TGGTGGGAT CTCTAGGAC CTAGTGGACA 1140  
 TCTGTATTA ATTTAATCTC AGGAAAAACA AGAAATTAAC CCAGAGAGAG TCTGGGTTTT 1200  
 GGAATTCAGC GTAGCTGACT CCAGACGCTG GTGCTGGCC TCCATTTTGT TCTGTCTTC 1260  
 25 AGCTCTGACT TACAGCTGCA GTCACTTTTG CTATAAGGCA CCGGGGTAGA AGGGTGGATG 1320  
 GGCTTCACAT CAATTTTTTT CTTCCTTTAG GGTGGGGGAT TGGTTTGGCT TTCTTTTGT 1380  
 GTGGTTTTTT GTTTTATTTT TGTCAAGATT GATTTTTAGA TGCAAGGACT TGAAGAAGACC 1440  
 CAGAAGATGT CCACCAAGTT TTCCCTGAGG CCTAGGATTT TTTATCTGT CCCGAGCAGA 1500  
 GGTAAATCTC CACACTTAG TGCAACAGTA GCACCAAGCA TTTTGAGCAG AGTACCTCTT 1560  
 30 TGGGAGGCTT TCGTTTGTG TTTGTTTTTA ATTCTCTTTC CTTAGCAGCA AGGTCTTTTT 1620  
 TCCTAGAGAA TCTACTCCGT TGCAGAAATC TTGCAACCTC AGGAGCCCTC ACTGATTGAG 1680  
 TGCTGTCAGC CTGATATACT ACTTGGACT CTGGAACAGC ATATGGGTTT TATCTCTAT 1740  
 TTCTACTGTG TGTGTTAAA CAACCGTCGG AGACCAAGAT ACCTGTTAGA TGGCTAGTCC 1800  
 TGTATAACTC GACTCTGTAT GTTTCAATGT ATGTTACTGC AATGCTTCAC CTGCTGTACA 1860  
 35 GTGTTGTGA GATGCTCTTT GAAGATGGTA CTTTATATT T

Seq ID NO: 232 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 40 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRF TRPLECQDAL ETAARAEGLS 60  
 LDASMSHQLR ILDEEHPKXK YHGLSALKP IRTTSKHQHP VDNAGLPSCM TFSWLSLAR 120  
 VAHKKGELSM EDVNSLSKHE SSDVNCRRLE RLWQEEINEV GPDAASLRV VWIFCRTRLI 180  
 45 LSIVCLMITQ LAGFSGPNFQ DGCILRSE

Seq ID NO: 233 DNA sequence  
 Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
 50 TTTTAATGGT GCTCATATAT ACTGTATTTT TTGTTGTTTA GTTTTACTTA TTGAGAGTGT 60  
 CACAACATGA ATCACAATAT CATGATTTTT TTTTCTTACT TTTACTCCCC AAATTATTCA 120  
 TGTTCCTTAG ATCGTAGTCA TTGAGAAGTC CCAATAACTC TAAACTTTTG AGTTATAACG 180  
 TAGTAACACT CTCTTTCATC TTTGTGTAG CTCTGTAGTC TTAACCTGGA TTTTAATTTT 240  
 55 TTTGTTTCCA AAGTCACAAT TGAATTATTC TTAGATACCT TAAGCCACTG AATTCAGTTC 300  
 TGTTTGACTG AAAGCAAAAAC AAGTGACAG TTTATTTTCA AACACTAART TCTTGATATT 360  
 TGTATTATGT ATATCTTTTT ATTAATATAT TATTTTGACT AAGCTTTCAT AAAATATTG 420  
 AAGCTATTTT AATCATCAAG TATGAAAAC AAATTACTAT TGCAATTTCC TATATATGCA 480  
 60 TATATTATGG ATTAACCGA ATGTATCAT TTTTGGCCTA ATGTCTGGAT ATAAAAGATA 540  
 ATTAGCTAC TATAGTATTA ATAAATTTT CAGTTGGTTT GGGCAAAATT AAACCTGAAA 600  
 AATAGTTAA AAGTAGTTA CAAATTAAC TTAATAATT ATACCTGATT TTTTCTCTG 660  
 AAATTAAGTA CATTTTAAAT GAGCTTTATA ATACCTTAA AAGTTGGTTC TAATTTAAAA 720  
 TATGAAAGCT CTGGCTATCA TCCTGGGATA GTAATTTCTA ATTATATAGT ATTTCAAAAC 780  
 65 TATATATTTT TTAGTTCCTT TGAGATAACT AATTTCTAAT TATATATGTT TCAAAAACCA 840  
 TATCCTGTAT TTTTCTTAAG AATTGTTTTA TAAATAGGTC ATAAGATACA AGGTCTGCAT 900  
 TAGAAGACCC ACTCTTACTA GGTTCCTTAA GGATCTGCCA TAGATTTTTT TTTTCTTTT 960  
 TTTTCTTTAG GTAGTTTAAA GCAAGCACTG ATACCACTGG GAGTTGGTCT TGATCTAGGA 1020  
 70 GATTCGTGTA AGCATCCAAA AACAAATGCT AATTTAGTCT CTTAGGTTAT GGCCTGTGAC 1080  
 TCCAGATAAA AGATGGAGAA TACCTCATGT ACTGTGACTT GAAATGAAT TCTTAAATTT 1140  
 CTTAGGCTCT CTCCATGTAT CTTTCTTAA GAAAAGTTTC TGAGTGTGAT CTCTCTTTG 1200  
 CCATAGTATC AAGTGGAGGG TAGTTCAGAA AAGTTAATAG GAAATCTTT GTGACAGCAG 1260  
 ACTATAATAG AAGTTTGAGT AATATTTTAA TAAATTTATA TAATTCAAAT GATAAAAATG 1320  
 75 TATCAATGTT ATCCAATGAT TTTTATTAAA AAATTACCTT ATTATTAGAA CTGTGCCTAT 1380  
 TACATAAAAA GTGCTCATGT ATTGAAATTT TAAATAATTT ATTTAAATCA AGACCACCAT 1440  
 AAGTCATTAA TAATTTAATA ATTGTTTTAA ATCAGTGGTT TTCAACCTTC ACTTCATATT 1500  
 AGAATCATCT GAGGACTTTT AATATGGAAT CCACCTCATA ACAATTAAGT CTAATTTCT 1560  
 GGAAGATGGA GCCATGCTTG TTTTCCAAA AGCTCTTGA GTGATCTAA TTTGTAGTCA 1620  
 80 GAGTTGAAGA CCAGTCTCTT AAATTAGTGC AGGAAAATGC TTTTATTCTT CCCATGTTAA 1680  
 CTTTAAAACT TAGTAATGTA CCCAGTTAAG TTTTGAATGT TTAATTTCCA CTAAAGAACA 1740  
 TATTTCTTCA ATAACTAGCA TTTATTACAT GAAATTTAAG AGTTTAAAGT CCATCAAACT 1800  
 AGCCCTGTGT TAAGATTATT ATTTCTTCTC TATAACTTCA AAATAGATAT TTTCTTCAA 1860  
 CTGTTGAGGT GAGAAAACAT AATGATTTT TTTTCTTTC CTCTGGAGCT GCCTGTTCAG 1920  
 85 TGAGATGGAG GAGGTGGGCA CATTTAAGGT CAGTTCACCT ACCATGGTGT CAGAGTTCTG 1980  
 ATCATATGGA AGTTTGGAAA AGAGAGCTTA TCACAGGTTT GTATGCTGGT GAATGGATAG 2040  
 TTTTAATTTCT CACTGTCTCA AAGAGAATC AGCTCTCCAG CAGTTCTAGA AAAGCTTTGA 2100  
 CAATCCCCAA GGGCAGTGT TACCTTACTC CTTCATGCT TCTTAGAAGG TAGAATTAAG 2160  
 TTTCTGGAAT TGCACTTACA GTTTTCTTA TTAACATTCA GAATTGGGAA TATTAATTTT 2220

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

TCCAGTGAAGT AGTTTTCTGA AATTGGTAAC TTGGAGAGTA AAATAACGTA TTTTGCTTTT 2280  
CAATTTTGTG TTGTTTACT TTTATGTAAA AATTTGATAT GTGAATTACA CAGTTCTAAT 2340  
AAAACCTCAT GCCTTTTCAT TACATCTAAT TTGAACTCTC AACTTCAGTG CCAGAAGTGC 2400  
TTTAAAGATG CTTTAAATGA AAGTATTAAG AAAATATATA GATTTGTATG TCAGTTTATA 2460  
CTTCAGAAAT CCATATATTT GTCATATTTA TTTTITTAGA AACCTCCTAA TTGGATAACT 2520  
AGATGGTATT TAAATGAAT GCCCAAAAAT ATCTTGTAAC TTTGTCCAAA AGTTTATCTG 2580  
TTGGAAGCCG CCAGCCATTC ATGTAGAGAG TTTATAAGAA AATAATTTAA AATTGTATGC 2640  
ATTTATATAT ACTATGGTAT CTGTGTACCA TATTCTAAG TATTCATTAT TAAATTGGTA 2700  
CTTCTTAAAA CCATACCTG GCTTGCCITT TAGTGTAAA CACAAAATCC AACATTGTAT 2760  
ATAGAGATTG TTCTTTTATG AAGAAGAGCT GACGTAATTT ATTACCAGTG CATCTGCACA 2820  
AAGACATTAA CATAAGTCTC TGAGCAGTGA TACATTTTCA AACATGAAGA GTGACAAACA 2880  
CCACATTAAA CRACCACGGC AACACTCAGA CTTGCACTT TCCTACGAAT CCATCCTATA 2940  
TGTGCTCGT ATCGCTCTG GCATAACTTA CACGAATCGT CCTCCCTACT TGTCTAGCT 3000  
CCTTCATCAA GCACCTGCCA ACACATTCAC CTCTAACTTG TACAACCTTA CCAACTCACC 3060  
ACAACATCTG CAACTCTACC CTATCAACTG CCAACTTAAA GACCCCAAC ACAACACAAC 3120  
CCCCAAACAC AAAACCACTA AATCATAACC ACCACACAC CCACACACCA CACACCCACC 3180  
CACACAAACA ACACACACG ACCAAACACC CCACCAACAA CAGCTAACA ACCACAAACA 3240  
GACAACACAT CACATACACT CACTACCCCC CCATCTCCC ACCCAACA

Seq ID NO: 234 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 27-281

1 11 21 31 41 51  
AGCAGGAGGA GAGCTGGCGG GAAGACATGC ACCCTTGAA GACCCAGAGA GAGGCGGTCT 60  
GTCTACCGCG TAGCAGTTAC ATCAGACTGA GACACTTCCT GTTTACAGGA GACTATAAAA 120  
TTCTGCCCC GTGCTCATTT GGGGCTGACG CCATTTTAGG CCTCAGCCCA TCTGCACCCA 180  
GGCGCTCACT GAAACAGTGT GTTGCTCCAC ACCGCTTGT TTTGCTTGT GGCGCGCTCT 240  
CAGGGTTCGG ACCAATCCAA GAGCCTTGCA GAAAGCATT ACGTGCTTTT CTCTTTGGCA 300  
GAGTTTTTCT TTGCTCTGAT CTTGGAGACA TCCCTCTGCC TAGTGGAAAC ATAAGGAATA 360  
CAGAAAGAAAT GCAAGGAGAT AGACCAACGT GAGATTCTCC TTCATGCACT CAAGAGAAAG 420  
ATGTTGACAG AAGAGCTAGT CTTTCAGGCT GGGCTGGTGA CCTGAGAAAG AATGTCCAGC 480  
TTTTCTTCTC CACTTGGCAT ATCAAGAGCC AGGCGTGGAA GACTAAAAA GGAATGTTT 540  
ATAAAAACTG TTCAGCGTGT CGCCAAACAG AAGTGGTAAA GTAGCAAAAA TGGGGATGGA 600  
GATGCCAGGA GAAAGATGTC CAGGGGTAAA GTGGGAAAT GGGAACTGA AGCCAGGAGG 660  
TCAGGCCAAG CCAACAGGTG TTCTGTTTTT CATCACAGAA CTAATAAGTG GTGCTGAGGA 720  
CTCAACCCG GGAAGGCCA CTCTGAACC CATGCTGGTC ATCCATATCC CCAAGGCCCT 780  
GGTCAGAAAC CAGCTAAGCA GATGGCTTGG GTCATCAGGA CGTCCATTAC ATCCAAAGGA 840  
AGACAGCCTG TGACGTTTCA AAAGCAAAAG TCCCTTACCA GCCAGTGAAG CTACCTGATT 900  
TCTCAGTATC TTACGCCGAG TGACACGATC TACCTCAAA ACTTAAAAA AAAAGGGAAA 960  
CATAAACACA TAACAGCAGC AGCAATAATT AAAGATGAGA TGAGAACAA TAAGAAAAA 1020  
GGAAAGGTCT CCTGTGACTG TTTTATTTT AGGGAACAG AGAGGAAGAA GAATGATTTT 1080  
TCTTTTGATG ACTCTATATC CAACTCTGAG GTTTGATTAA AGAAATGACC TTGAACCACA 1140  
GCAAGAAAAA ATAAAGACA ATTTCCAGTA AGTATGCCAG TTCGAATTAA TGATTTACTT 1200  
TTTATTTTAA AACCTGAATC AGCAGAGATT TACATGCATT ACGATGATTA ACATCTGAAA 1260  
TTTGACCTTG AATAATCTT TACATTGTAA ATTCTTAATG ATCAAAACAA GGTTCTCAGT 1320  
GATTAAACCA TATTAGTAA TAATTATTAA AGGAGAATAA TTGCAATAC AACATTCTTA 1380  
AAATCTCAAG GCTTTTAAAG CATTGTGACA AATGACTGGA CATTTTTAA ATTGAAAAA 1440  
AAAAAAAGC CCTCCATCTG ATTCTCAATT TCATTGTGAG TGCAACAACA AAAAAGGTAT 1500  
GCATTCTCT TCTCATTTTC CACTGTCTCG CAGCTAGAAA ATTCTCACGA CTACCTTTGA 1560  
TCCATCAAAA GCCAAAGAAA GAAAGAAAAA TTGTCTGTGA CAGATATATG ACATTAAAAA 1620  
ATAATCCC

Seq ID NO: 235 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
MHPLKQREA VCLPRSSYIR LRHFLFTGDY KIPAPCSFGA DAILGLSPSA PRSLKQCV 60  
PHRLVLLVGA LSGFRPIQEP CRKH

Seq ID NO: 236 DNA sequence  
Nucleic Acid Accession #: NM\_002075  
Coding sequence: 406..1428

1 11 21 31 41 51  
CCCAATAGG GGCAGACCTG TCCATCCTTC TCTGTGGGTC CCCTGTACCT TTCTCCCCCA 60  
ACAGGATCAG ACCCAGAGGC AGCTGGTTGG GGTITGTCGA GAAGAAGGAT TATCCAGATC 120  
AGTCTTTTCT AATCTCAGCT CTTGCTGTA CCTCCCATTA CTCACCAAC CCTCTCCCC 180  
ACCACCTTGA GCTGAGGAGC ACAGTTTGAG GCCCCCCCAA CCCCCCGCGG GTGCGGGCCA 240  
GGCCAGGCCA GGCAGCTCC TCTGGCAGCA GAGCCTGGGC AGGTGACGGG CGGGCGCGGG 300  
CGTCGAGCT GAGGAGTAA GGAGGCTCCC AGGAACCGGA GCTGGAACCC CGGCCGAGGT 360  
CCAGCCAGAG CCAAGAGGCC AGAGTGACCC CTCGACTGCT CAGCCATGGG GGAGATGGAG 420  
CAACTGCGTC AGGAAGCCGA GCAGCTCAAG AAGCAGATTG CAGATGCCAG GAAAGCCTGT 480  
GCTGACGTTA CTCTGGCAGA GCTGGTGTCT GGCCTAGAGG TGGTGGGACG AGTCCAGATG 540  
CGGACGCGGC GGAGCTTAA GGGACACCTG GCCAAGATT ACGCCATGCA CTGGGCCACT 600  
GATTCTAAGC TGTCTGTAAG TGCTCGCAA GATGGGAAGC TGATCGTGTG GGACAGCTAC 660  
ACCACCAACA AGGTGCACGC CATCCCACTG CGCTCCTCCT GGGTCATGAC CTGTGCTAT 720  
GCCCATCAG GGAACCTTGT GGCAATGGG GGGCTGGACA ACATGTGTTT CATCTACAAC 780  
CTCAATCCC GTGAGGGCAA TGTCAGGTG AGCCGGGAGC TTTCTGCTCA CACAGTTAT 840  
CTCTCCTGCT GCGGCTTCTT GGATGACAAC AATATTGTGA CCAGCTCGGG GGACACACG 900  
TGTGCTTGT GGGACATTGA GACTGGCAG CAGAAGACTG TATTGTGGG ACACACGGGT 960  
GACTGATGA GCTCGGCTGT GTCTCCTGAC TTCAATCTCT TCATTTCGGG GGCCTGTGAT 1020  
GCCAGTGCCA AGCTCTGGGA TGTGCGAGAG GGGACCTGCC GTCAGACTTT CACTGGCCAC 1080

5 GAGTGGGACA TCAACGCCAT CTGTTTCTTC CCCAATGGAG AGGCCATCTG CACGGGCTCG 1140  
 GATGACGCTT CCTGCGCTT GTTTGACCTG CGGGCAGACC AGGAGCTGAT CTGCTTCTCC 1200  
 CACGAGAGCA TCATCTGGCG CATCACGTCC GTGGCCTTCT CCTCAGTGG CCGCTACTA 1260  
 TTCCTGGCT CTGGCCACGA TAACAGGGTG AGCTGCCTGG GAGTCACAGC TGACGGGTG 1320  
 GGCACTCTCT CTGGCCACGA TAACAGGGTG AGCTGCCTGG GAGTCACAGC TGACGGGTG 1380  
 GCTGTGGCCA CAGGTTCCCT GGACAGCTTC CTCAAAATCT GGAAGTGGAG AGCTGGGAGA 1440  
 AAGGGAAGTG GAAGGCACTG AACACACTCA GCAGCCCCCT GCCCGACCCC ATCTCATTC 1500  
 GGTGTCTCT TCTATATTC GGGTGCCATT CCCACTAAGC TTTCTCTTT GAGGGCAGTG 1560  
 10 GGGAGCATGG GACTGTGGCT TTGGGAGGCA GCATCAGGGA CACAGGGGCA AAGAACTGCC 1620  
 CCATCTCTCT CCATGGCCTT CCTCCCCAC AGTCTCTACA GCCTCTCCT TAATGAGCAA 1680  
 GGACAACCTG CCCCTCCCCA GCCCTTTGCA GGCCACGAG ACTTGAGTCT GAGGCCCCAG 1740  
 GCCCTAGGAT TCCTCCCCA GAGCCACTAC CTTTGTCCAG GCCTGGGTGG TATAGGGCGT 1800  
 TTGGCCCTGT GACTATGGCT CTGGCACCAC TAGGGTCTGT GCCCTCTTCT TATTCATGCT 1860  
 15 TTCTCTTTT TCTACCTTTT TTCTCTCTCT AAGACACCTG CAATAAAGTG TAGCACCTCG 1920  
 GT

Seq ID NO: 237 Protein sequence:  
 Protein Accession #: NP\_002066

20 1 11 21 31 41 51  
 MEEMBLRQE AEQLKKQIAD ARKACADVTL AELVSGLEV GRVQMRTRT LRHLAKIYA 60  
 MHWATDSKLL VSASQDGLI VWDSTYTNKV HAIPLRSSNW MTCAYAPSGN FVACCGLDNM 120  
 CSIYNLKSRE GNVKVSRELS AHTGYLSCCR FLDDNNIVTS SGDITCALWD IETGQKTVF 180  
 25 VGHGTGDCMSL AVSPFNFLEI SGACDASAKL WDVREGTCRQ TPTGHESDIN AICFPFNGEA 240  
 ICTGSDDASC RLFDLRADQE LICFSHESII CGITSVAFSL SGRLLFAGYD DFNQNVWDSM 300  
 KSERVGLISG HDNRVSLGLV TADGMAVATG SWDSPLKIWN

Seq ID NO: 238 DNA sequence  
 Nucleic Acid Accession #: CAT cluster

30 1 11 21 31 41 51  
 TCCCAATGTG TNGAACCTAC CATAAATCTT TTCTTACNG GACAATCTTA TNCTAANCAA 60  
 35 TACCATTGTC TTTTAAAGCA GATAATCCTC CAAGTTTCT AATGATATCT GAAACTATTA 120  
 ACTGATTCTG TGAATTATGA AATCTGAAAA GGAATTGGAA GTTGCTAAAA ATCTATCATT 180  
 TGCAATTGACC AGTGTGAAGC ACAGTGGAAAT GAGAATGGCT GCCTGACAC CAAAGAAAAA 240  
 TAAGTGACTG GAAAGCTGAA GAATCACCGG CTTCACTGAC ATGGAACCCA GTGATTGAT 300  
 40 TTTTGACGAG TATCGGGTGA CTTTGAGGTG GTCAAGAAAC CACACTTTAA GAACAATGTC 360  
 CAAAAGGGG AAAAAAAGA GCAACCAAG AAAAAAATC CATAAATTTG CACAGAAGAA 420  
 AAGAAAGAAA AATAAATAT ACAATATGGA CGATGGAGAA AAACAGTTAC ATTTCTTTAT 480  
 GGATCAAGAA GTTTGTGTAC ACATAATCTC ATTTTGAGAT ATATACTAT TTTGTCTTT 540  
 CAGAAGTGA TCAAAATATT TCAAAATGCT GTCTTATGAA ACTACAATAT TCTCACAGAT 600  
 TAGAAAAATT TTCTGTAAA AGTCAGATAG TAAATATTTT AGGTTTGTGA GTGTCTTTTG 660  
 45 CAACTACTCA ACTTCTCTAC TGTAGCACAA GAGTAGCTGT GGTACTGTGC AAATAAATTG 720  
 CTTGTGTTCC AATAAAGCTT CATTACAAA AACATGCCAT GGGCCATATT TGGCCTGTAC 780  
 ACTGTGTTT GCCAAGTCTT AATATAGTTG CTTAGCAAGT ATTGTGAGCT ATTTGAGGAA 840  
 GAGATGAAAG TTCATTGGGT TGCTAAAAAG TATGTAGAAA TTCAAGGAA AATTAATAAT 900  
 50 TAGGCTAAGT TATAATACAC TGTTTTAAAC ATTGTAAAT GTAAGAGAAA TTTACAATA 960  
 AAAATCCCAA ATAAAA

Seq ID NO: 239 DNA sequence  
 Nucleic Acid Accession #: NM\_001786.1  
 Coding sequence: 130-1023

55 1 11 21 31 41 51  
 GGGGGGGGGG GGCACCTGGC TTCAAAGCTG GCTCTGGAA ATTGAGCGGA GAGCGACGG 60  
 60 GTTGTGTAG CTGCCGCTGC GGCCGCGCGC GAATAATAAG CCGGATCTA CCATACCCAT 120  
 TGACTAACTA TGGAGATTA TACCAAAATA GAGAAAATTG GAGAAGGTAC CTATGGAGTT 180  
 GTGTATAAGG GTAGACACAA AACTCAGGT CAAGTGGTAG CCATGAAAAA AATCAGACTA 240  
 GAAAGTGAAG AGGAAGGGGT TCCTAGTACT GCAATTCGGG AAATTTCTCT ATTAAGGAA 300  
 CTTCTGATC CAAATATAGT CAGTCTCAG GATGTGCTTA TGCAGGATTC CAGGTTATAT 360  
 65 CTATCTTTG AGTTTCTTTC CATGGATCTG AAGAAATCT TGGATTCTAT CCTCCTGGT 420  
 CAGTACATGG ATTCTTCACT GTTTAAGAGT TATTTATACC AAATCCTACA GGGGATTGTG 480  
 TTTTGTCACT CTAGAAGAGT TCTTACAGA GACTTAAAC CTCAAAATCT CTGATTGAT 540  
 GACAAGGAA CAATTAACCT GGCTGATTTT GGCCTTGCCA GAGCTTTTGG AATACCTATC 600  
 AGAGTATATA CACATGAGGT AGTAACACTC TGGTACAGAT CTCAGAGAT ATTGCTGGGG 660  
 70 TCAGCTCGTT ACTCAACTCC AGTTGACATT TGGAGTATAG GCACCATATT TGCTGAACCTA 720  
 GCAACTAAGA AACCACTTTT CCATGGGGAT TCAGAAAATTG ATCAACTCTT CAGGATTTC 780  
 AGAGCTTTGG GCACTCCCAA TAATGAAGTG TGGCCAGAAG TGGAACTTTT ACAGGACTAT 840  
 AAGAATACAT TTCCCAATG GAAACCAAGG AGCTAGCAT CCCATGTCAA AAACCTTGGAT 900  
 GAAATGGCT TGGATTGTCT CTCGAAAATG TTAATCTATG ATCCAGCCAA ACCGAATTTCT 960  
 75 GGCAAAATGG CACTGAATCA TCCATATTTT AATGATTGAG ACAAATCAGT TAAGAAGATG 1020  
 TAGCTTTCTG AAAAAAGTT TCCATATGTT ATGTCAACAG ATAGTTGTGT TTTATTGTT 1080  
 AACTCTGTG TATTTTGTG TATATATAT TCTTTGTTA TCAAACTCA GCTGTACTTC 1140  
 GTCTCTAAT TTCAAAAATA TAACTAAAA ATGTAAATAT TCTATATGAA TTTAAATATA 1200  
 80 ATTCTGTAAA TGTGAAAAA AAAAAA AAAA

Seq ID NO: 240 Protein sequence:  
 Protein Accession #: NP\_001777.1

85 1 11 21 31 41 51  
 MEDYTKIEKI GEGTYGVVYK GRHKTGQVQV AMKKIRLESE EGVVPSTAIR EISLLKELRH 60  
 PNIVSLQDVL MQDSRLYLIF EPLSMDLKKY LDSIPPQYIM DSSLVKSILY QILQGIVFCH 120

SRRLVLRDLK PQNLLIDDKG TIKLADFLGA RAFGIPRIVY THEVVTLYWR SPEVLLGSAR 180  
 YSTPVDIWSI GTIFAEIATK KPLPHGDSEI DQLFRIFRAL GTPNNEVWPE VESLQDYKNT 240  
 FPKWFKGSLA SHVKNLDENG LDLLSKMLIY DPAKRISGKM ALNHPYPNDL DNQIKKM

5

Seq ID NO: 241 DNA sequence  
 Nucleic Acid Accession #: NM\_033379.1  
 Coding sequence: 132-854

10 1 11 21 31 41 51  
 CGCCCGCGCG CGGGCTCAAC TTGTAGAGC GAGGGGCCAA CTTGGCAGAG CGCGCGGCCA 60  
 GCITTCGAGA GAGCGCCCTC CAGGGACTAT GCGTGCGGGG ACACGGGATC TACCCATACC 120  
 15 ATTGACTAAC TATGGAAGAT TATACCAAAA TAGAGAAAAT TGGAGAAGGT ACCTATGGAG 180  
 TTGTGTATAA GGGTAGACAC AAAACTACAG GTCAAGTGGT AGCCATGAAA AAATCAGAC 240  
 TAGAAGATGA AGAGGAAGGG GTTCTAGTA CTGCAATTCT GGAAATTCT CTATTAAAGG 300  
 AACTTCGTCA TCCAAATATA GTCAGTCTTC AGGATGTGCT TATGCAGGAT TCCAGTTAT 360  
 ATCTCATCTT TGAGTTCTCT TCCATGGATC TGAAGAAATA CTGGATTCT ATCCCTCCTG 420  
 20 GTCAGTACAT GGAATCTTCA CTGTAAAGG TAGTAACACT CTGGTACAGA TCTCCAGAAG 480  
 TATTGCTGGG GTCAGCTCGT TACTCAACTC CAGTTGACAT TTGGAGTATA GGCACCATAT 540  
 TTGCTGAAC AGCAACTAAG AAACCACTTT TCCATGGGGA TTCAGAAATT GATCAACTCT 600  
 TCAGGATTTT CAGAGCTTTG GGCACCTCCA ATAATGAAGT GTGGCCAGAA GTGGAATCTT 660  
 TACAGGACTA TAAGAAATACA TTTCCTCAAT GGAACACAGG AAGCCTAGCA TCCCATGTCA 720  
 25 AAAACTTGGG TAAGAAATGGG TTGGATTTCG TCTCGAAAAT GTTAATCTAT GATCCAGCCA 780  
 AACGAATTTT TGGCAAAATG GCACTGAATC ATCCATATTT TAATGATTG GACAATCAGA 840  
 TTAAGAAGAT TGAGCTTTCT GACAAAAGT TTCCATATGT TATGTCAACA GATAGTTGTG 900  
 TTTTATATGT TAACCTCTGT CTATTTTGT CTATATATA TTTCTTTGT ATCAAACTTC 960  
 AGCTGTACTT CGTCTCTAA TTTCAAAAAT ATAACCTAAA AATGTAAATA TTCTATATGA 1020  
 30 ATTTAAATAT AATTCTGTAA ATGTGAAAAA AAAAAAATA AAAAAA

Seq ID NO: 242 Protein sequence:  
 Protein Accession #: NP\_203698.1

35 1 11 21 31 41 51  
 MEDYTKIEKI GEGTYGVVYK GRHKTIGQV AMKKIRLESE EGVVPSTAIR EISLLKELRH 60  
 PNIVSLQDVL MDSRLYLIF EPLSMDLKKY LDSIPPQYIM DSSLVKVVTI WYRSPVLLG 120  
 SARYSTPVDI WSGITFAEL ATKPLPHGD SEIDQLFRIP RALGTPNNEV WPEVESLQDY 180  
 40 KNTFPKWKPG SLASHVKNLD ENGLDLLSKM LIYDPAKRIS GKMAIHPYP NDLDNQIKKM

Seq ID NO: 243 DNA sequence  
 Nucleic Acid Accession #: AF101051.1  
 Coding sequence: 221-856

45 1 11 21 31 41 51  
 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGCGCGGCC CGGGCGCGCG ACCCCCAACC 60  
 CGACCCAGAG CTTTCCAGGC GCGCGCGCAG CGAGCAGGGC TCCCGCGCTT AACTTCCTCC 120  
 50 GCGGGGCCCA GCCACCTTCG GAGTCCGGG TTGCCACCTC GCAAACTCTC CGCCTTCTGC 180  
 ACCTGCCACC CTTGAGCCAG CGCGCGCGCC CGAGCGAGTC ATGGCCAAAC CGGGCTGCA 240  
 GCTGTGCGG CTTGATCTCG CTTCTCTGGG ATGGATCGGC GCCATCGTCA GCACTGCCTC 300  
 CGCCCACTGG AGGATTTACT CCTATGCGCG CGACAACATC GTGACCGCCC AGGCCATGTA 360  
 CGAGGGGCTG TGGATGTCTG GCGTGTGCGA GAGCACCGGG CAGATCCAGT GCAAAGTCTT 420  
 55 TGACTCTCTG CTGAATCTGA GCAGCACATT GCAAGCAACC CGTGCCTTGA TGGTGGTTGG 480  
 CATCCCTCTG GAGGTCTCTG CAATCTTTGT GGCCACCGTT GGCATGAAGT GTATGAAGTG 540  
 CTTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTTCT 600  
 TCTTGAGGT CTGGCTATTT TAGTTGCCAC AGCATGGTAT GGCATAGAAA TCGTTCAAGA 660  
 ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTTGGTCAGG CTCTCTTAC 720  
 60 TGGCTGGGCT GCTGCTTCTC TCTGCTTCTT GGGAGGTGCC CTACTTTGCT GTTCTCTGCC 780  
 CGAATAAACA ACCTCTTACC CAACCAAGG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840  
 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAATCA TGTGAAACA AACCGAAAT 900  
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960  
 65 GTATGGTATT ACAAAACAAA CAACCAAGG AAAACCCAT GTGTTAAAT ACTCAGTGCT 1020  
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080  
 TTGTATTACT GCTTCCCATT GAGTAATCAT ACTCAAATGG GGGAGGGGT GCTCCTTAAA 1140  
 TATATATAGA TATGATATA TACATGTTT TCTATTAATA ATAGACAGTA AAATACTATT 1200  
 CTCATTATGT TGATCTAGC ATACTTAAA TATCTCTAAA ATAGGTAAAT GTATTTAATT 1260  
 70 CCATATTGAT GAAGATGTTT ATTGGTATAT TTTCTTTTC GTCTTATAT ACATATGTAA 1320  
 CAGTCAAAAT TCATTACTC TTCTTCAAT GCTTTGGGTG CCTTTGCCAC AAGACCTAGC 1380  
 CTAATTTACC AAGGATGAAT TCTTTCAATT CTTATGCGT GCCCTTTTCA TATACTTATT 1440  
 TTATTTTATA CCATAATCTT ATAGCACTTG CATGTTATT AAGCCCTTAT TTGTTTGTG 1500  
 TTTTATGGT CTCTATCTCC TGAATCTAAC ACATTTCTAT GCCTACATTT TAGTTTCTAA 1560  
 75 AGCCAAAGAG AATTATTATC AAATCAGAAC TTTGGAGGCA AATCTTTCTG CATGACCAAA 1620  
 GTGATAAATT CCTGTGACC TTCCACACA ATCCCTGTAC TCTGACCCAT AGCACTCTTG 1680  
 TTTGCTTTGA AAATATTGT CCAATTGAGT AGCTGCATGC TGTTCOCCA GGTGTTGTA 1740  
 CACAACITTA TTGATTGAAT TTTTAAGCTA CTTATTCATA GTTTTATATC CCCCTAACT 1800  
 80 ACCTTTTGT TCCCATCTC TTAATTGTAT TGTTTTCCA AGTGTAATTA TCATGCGTTT 1860  
 TATATCTTCC TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGCTAGAC TTTCTGGAGT 1920  
 GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAGCA AGTCACTTAA TCTTTCTACC 1980  
 TCTTTTCTT ATCTGCAAAA TTGAGATAAT GATACCTAAC CAGTTAGAAG AGGTAGTGTG 2040  
 AATATTAAAT AGTTTATATT ACTCTCATTC TTGAACATG AACTATGCCT ATGTAGTGTG 2100  
 TTTATTTGCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACCT ACACAGTAC 2160  
 85 CTTATATGTA TTAATCTCTC TCTCTCTCT ACCAGTCTAT TTCCACTGAA CAAAACCTAC 2220  
 ACACATACCT TCATGTGGTT CAGTGCCTTC CTCTCTCTAC CAGTCTATTT CCACTGAACA 2280  
 AAACCTAGC ACATACCTTC ATGTGGCTCA GTGCTTCTCT CTCTCTACCA GTCTATTTC 2340  
 ATTCTTTCAG CTGTGCTCTG CATGTTTGTG CTCTGTTCCT TTTTAAACAC TGCTCTTACT 2400  
 TTTCCAGTCT GTACAGAAAT CTATTTCACT TGAGCAAGAT GATGTATGGA AAGGGTGTG 2460



5 GCACTGGTGT CTGGAGACCT GGATTGAGT CTGGTGCTA TCAATCACCG TCTGTGTTTG 2520  
AGCAAGGCAT TTGGCTGCTG TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAATTC 2580  
CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAAT ACATGTAAGT 2640  
GTGGTTTTGT AATTGTAAGA GTGCTATACT AAGGGAAAGA ATTGAGGAAT TAACTGCATA 2700  
CGTTTTGGTG TTGCTTTTCA AATGTTTGAA AATAAAAAAA TGTTAAGAAA TGGGTTTCTT 2760  
GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAAACGAAT 2820  
AAGATTCTGA GGAAGTCTTA TCTTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880  
ACAGATGTAA TGGGAAGAAA TAAAAGCCTA CGTGTGTGTA AATCCAACAG CAAGGGAGAT 2940  
TTTGAATCA TAATAACTCA TAAGGTGCTA TCTGTTCAAT GATGCCCTCA GAGCTCTTGC 3000  
TGTTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGGTACT TCATAATAAA 3060  
CTACACAAGG AAAGTCAGCC ACCGTGCTCT ATGAGGAATT GGACCTAATA AATTTTAGTG 3120  
TGCTTCCAA CACTGAGAAT ATATGCTTTT GGAAGTTAAA ATTTAAATGG CTTTGGCCAC 3180  
ATACATAGAT CTTCAATGAT TGTGAGTGTA ATTCATGTG GATATCAGTT ACCAAACATT 3240  
ACAAAAAAT TTTATGGCCC AAAATGACCA ACGAAATTGT TACAATAGAA TTTATCCAA 3300  
15 TTTGATCTTT TTATATTCTT CTACCACACC TGGAAACAGA CCAATAGACA TTTTGGGGTT 3360  
TTATAATGGG AATTGTGATA AAGCATTACT CTTTTTCAAT AAATTGTTTT TTAATTTAAA 3420  
AAAAGGAAAA AAAAAA AAA

20 Seq ID NO: 244 Protein sequence:  
Protein Accession #: AAD16433.1

25 1 11 21 31 41 51  
MANAGLQLLS FILAFLGWIG AIVSTALEFQW RIYSYAGDNI VTAQAMYEGE WMSCVVSQSTG 60  
QIQCKVFDLS LNLSSLTQAT RALMVVGILL GVIAIFVATV GKCKMKCLEL DEVQKMRMAV 120  
IGGAIFLLAG LAAILVATAWY GNRIVQEFYD PMTPVNARYE FGQALFTGWA AASLCLLGGA 180  
LLCCSCPRKT TSYPTPRPYP KPAPSSGKDY V

30 Seq ID NO: 245 DNA sequence  
Nucleic Acid Accession #: CAT cluster

35 1 11 21 31 41 51  
TTTTTTTTTT TTTTTTTTTT TTTTCAAGG AGAGCACAAG GAACCTTATT AATGACTTTC 60  
TTAATGGTTA AATGCTGTTT ACCAAGTGAC CCAGAGGCAG CGTGGTTTAG TGGTTTCAAC 120  
AGCATGGTCC CGAGAGTCTG ACAAACCTCA GTTCAATACC TTCTTTTGTC TTCACTTAGT 180  
TTTTCTTCTT GAGATTTAGT TTCTTCTATG TTAACAATGA GGATATTAAAT ATGTTTCACA 240  
40 CAGTTGTTAT GAAGAATGCA TATATTAGAA TGCTGTAGT CTCAGCTACT CAGGAGGCTA 300  
AGGTGGGGAG GTGCTCAAG CCCAGGAATT CAAAGCTGCA ATGCATTATG ATTACAGCTG 360  
TTAATAGCCA TGCACTTCA GCCTGGGCAA TGTAGTAAGA TCCCATCTCT GCCTCGGAGG 420  
GTCTACGCC CACGGAGTCT CGCTGATTGC TAGCACAGCA GTCTGAGATC AAACGTCA

45 Seq ID NO: 246 DNA sequence  
Nucleic Acid Accession #: XM\_058553.2  
Coding sequence: 897-1400

50 1 11 21 31 41 51  
AATTTTCAGA AGTTTCGTAT GGGGATGGTT TTATATAAAT TCAGGTTTTT CCCACAATAA 60  
TAAATGTATT TAGTCTCAGT GCTCAATAGA AGAGATTTC T AATAGAAAAG GATTCAAAC 120  
GTGAAACCAT TTCTCTTTTA ATGTTTACA TTCTGTGTAC AGATTGTTC TCTTGTGACT 180  
CTGTTATCCA TAATATGGAC AGTTCTTGAG TCCTAACATT GAGAGGTTTT CCCTTAGTGC 240  
55 ATAGAGGGAA TGAGTATTAA TTGGAGAAGC TTAAGTATT GCCACTTTAG CACTGAAGAT 300  
TGGGATGAGA GGAGGTGAAA CCTCACTAGA AAAAGGGACA ATGTTAGTGT GGCCCTTCCT 360  
GATCATGTTT AAGAAAAGTC ATGAAAATGG TGAAGTAGTG TTTCCAAGCA TATTGGAAGG 420  
GTTGAGTGTA TACTGTCTGT CAAAGACTTC CAGCATTTC AGTCTCTAGA GAGGAACAAG 480  
ACTGTAAACC TGCTTATCTG TATTTTAAAG AACCACAGGAG GAAAGCTTTA TAATAGAAC 540  
60 TTATTTCTGT GTTTATGTAT AAGGGGTTTT TTGTTTTTTT AAAGACAGGA TCTCACTCCA 600  
TTGTCAGGC CAAGTGCAAT GGCACGAACC TCATAGCTCC TGGACTTAAG TGAATCTGCT 660  
GCCTTTGCCT CTGAGTAGC TGGGACTACA GGCATGAGCC CCCATGCTG GCTAAGTTTG 720  
TTTTTTTGTG TGTGTTTGTG TTTGTTTTTG GGGGGGGTTG TTTGTTTTTT TGTAGAGACG 780  
TAGTCTTGCT TTGTTGCCAG GCTAGTCTCA AACTCTGCGC TTCAGTGAT CCTCCTGCCT 840  
65 CAGCTCCCA GAGTGCTAGG ATTACAGCAC TTGGATTGAG CTTCTTCAAT TCCAACATGG 900  
AAGAACTTA CACCGACTCC CTGGACCTCG AGAAGCTATT GCAATGCCCC TATGACAAA 960  
ACCATCAAAT CAGGGCTTGC AGGTTTCCCT ATCATCTTAT CAAGTGACAG AAGAATCATC 1020  
CTGATGTTGC AAGCAAAATG GCTACTGTGC CCTTCAATGC TGGCCACAG GTTCTCTGAG 1080  
CTGAAATTAG TCATCATATC TCAAGCTGTG ATGACAGAAG TTGATTGAG CAAGATGTTG 1140  
70 TCAACCAAC CAGGAGCCTT AGACAAGAGA CTCTGGCTGA GAGCACTTGG CAGTGCCTTC 1200  
CTTGGATGA AGACTGGGAT AAAGATTGTG GGGAGCAGAC CAGCACCCCA TTTGTCTGGG 1260  
GCACAACCTCA CTACTCTGAC AACACAGCC CTGCGAGCAA CATAGTTACA GAACATAAGA 1320  
ATAACCTGGC TTCAGGCATG CGAGTTCCCA AATCTCTGCC GTATGTTCTG CCATGGAAAA 1380  
ACAAATGAAA TGCACAGTAA CTGAATACCT ATCTCATCAA ATGCCAGACC CTAGAAGACT 1440  
75 GTTGCTCTT CTTTACAGG TGGGTTCTCA TTTTCTCTCT AATCTAATTA TAGAATGGTA 1500  
AACTCCCTGT GACTTTCCAA ACTGACAAGC ACACTTTTTT CCTCCCCCT TGAATCTCTA 1560  
TTTATGCAA GAACCTCAT ACTCAGAAGC TTCCAAATAA ACCTTTGATA CAGATTG

80 Seq ID NO: 247 Protein sequence:  
Protein Accession #: XP\_058553.1

85 1 11 21 31 41 51  
MEETYSLSL PEKLLQCPYD KNHQRACRF PYHLIKCRKN HPDVASKLAT CFPNARHQVP 60  
RAEISHHLS CDRSCIEQD VVNQTRSLRQ ETLARSTWQC PPCDEDWDKD LWEQTSFPFV 120  
WGTTHYSDNN SPASNIVTEH KNNLASGMRV PKSLPYVLPW KNNQNAQ

Seq ID NO: 248 DNA sequence  
Nucleic Acid Accession #: NM\_003392  
Coding sequence: 758..1855

5  
1 11 21 31 41 51  
TTAAGGAAAT CCGGGCTGCT CTTCCCCATC TGGAAAGTGGC TTCCCCCACA TCGGCTCGTA 60  
AACTGATTAT GAAACATACG ATGTTAATTC GGAGCTGCAT TTCCCAGCTG GGCACCTCTCG 120  
CGCGCTGGTC CCGGGGGCCT CGCCCCCCAC CCCCTGCCCT TCCTCCCGGC GTCTGTGCCCC 180  
CATCCTCCAC CCCCGCGGCT GGGCACCCCG CCTCCTTGGC AGCCTCTGGC GGCAGCGCGC 240  
TCCACTCGCC TCCCGTGTCT CTCTGGCCCA TGGAAATTAAT TCTGGCTCCA CTGTGTGCTC 300  
GGCCCGAGTT GGGGAGAGGA CGGAGGTGGC CCGCAGCGGC TTCTGTAGTG AATTACCCAG 360  
GAGGGACTGA GCACAGCACC AACTAGAGAG GGGTCAGGGG GTGCGGACT CGAGCGAGCA 420  
GGAAGGAGCG AGCGCTTGGC ACCAGGGCTT TGAATCAACA GAATTGAGAC ACGTTTGTAA 480  
TCGCTGGCGT GCCCGCGCGA CAGGATCCCA GCGAAATCA GATTTCCTGG TGAGGTTCGG 540  
TGGGTGGATT AATTTGAAAA AAGAAACTGC CTATATCTTG CCATCAAAAA ACTCACGGAG 600  
GAGAAGCGCA GTCAATCAAC AGTAAACTTA AGAGACCCCG GATGCTCCCC TGGTTAACT 660  
TGTATGCTTG AAAATTATCT GAGAGGGAAT AAACATCTTT TCCTTCTTCC CTCTCCAGAA 720  
GTCCATTGGA ATATTAAAGC CAGGAGTTGC TTTGGGATG GCTGGAAGTG CAATGTCTTC 780  
CAAGTTCTTC TGCTGGCTT TGGCCATATT TTTCTCTTTC GCCCAGTTG TAATTGAAGC 840  
CAATCTTGG TGGTGGCTAG GTATGAATAA CCTGTTCAG ATGTGAGAAG TATATATTAT 900  
AGGAGCACAG CCTCTCTGCA GCCAATCTGG AGGACTTTCT CAAGGACAGA AGAAACTGTG 960  
CCACTTGTAT CAGGACCACT TGCAGTACAT CGGAGAGGCG GCGAAGACAG GCATCAAGA 1020  
ATGCGAGTAT CAATTCCGAC ATCGACGGTG GAACCTGCAG ACTGTGGATA ACACCTCTGT 1080  
TTTTGGCAGG GTGATGCAGA TAGGACGGCG CGAGACGGCC TTCACATACG CGGTGAGCGC 1140  
AGCAGGGGTG GTGAAGCCCA TGAGCGGGCG GTGCGCGAG GCGAGCTGT CCACTGTGCG 1200  
CTGACGCCGC GCCCGCGGCC CCAAGGACCT GCCCGGGGAC TGGCTCTGGG GCGGCTGCGG 1260  
GCACAACATC GACTATGGCT ACCGCTTTGC CAAGGAGTTC GTGGACGCC GCGAGCGGGA 1320  
CGGCATCCAC GCCAAGGGCT CCTACGAGAG TGCTCGCATC CTCTGTAACC TGCAACAACA 1380  
CGAGGCCGCG GCACAGGACG GTTACAACTT GGTGATGTG GCTGTCAAGT GCCATGGGGT 1440  
GTCCGGCTCA TGTAGCTGTA AGACATGCTG GCTGCAGCTG GCAGACTTCC GCAAGGTGGG 1500  
TGATGCCCTG AAGGAGAAGT ACGACAGCGC GCGCGCCATG CGGCTCAACA GCGCGGGCAA 1560  
GTTGGTACAG GTCAACAGCC GCTTCAACTC GCCCACCACA CAAGACCTGG TCTACATCGA 1620  
CCCCAGCCCT GACTACTGGG TGGCCTAATG GAGCACCAGC TCGCTGGGCA CGCAGGGCGG 1680  
CCTGTGCAAC AAGACCTCGG AGGGCATGGA TGGCTGGAG CTGATGTGCT GCGGCGGTGG 1740  
GTAGACCAAG TTCAAGACCG TGCAGACGGA GCGCTGCCAC TGCAAGTTCC ACTGGTGTG 1800  
CTAGCTCAAG TGCAGAAAGT GCACGGAGAT CGTGACCAG TTTGTGTGCA AGTAGTGGT 1860  
GCCACCCAGC ACTCAGCCCC GCTCCAGGA CCCGCTTATT TATAGAAAGT ACAGTGATT 1920  
TGGTTTTGG TTTTATGAAA TATTTTTTAT TTTTCCCCAA GAATTGCAAC CGGAACCAT 1980  
TTTTTCTCTG TTACCATCTA AGAATCTGT GGTATTATAT TAATATTATA ATATTATT 2040  
GGCAATAATG GGGGTGGGAA CCACGAAAAA TATTATTTT GTGGATCTTT GAAAGGTAA 2100  
TACAAGACTT CTTTTGGATA GTATAGAATG AAGGGGGA TAACACATAC CCTAATTAG 2160  
CTGTGTGGGA CATGTACAC ATCCAGAAGG TAAAGAAATA CATTTCCTTT TCTCAAATA 2220  
TGCCATCAT TGGGATGGGT AGGTTCCAGT TGAAGAGGG TGGTAGAAAT CTATTCAACA 2280  
TTAGCTTCT ATACCAAAA TGAGTTGTAA ATTCTCTGGT GCAAGATAAA AGGTCTGGG 2340  
AAAAACAAAC AAAACAAAC AAACCTCCCT TCCCCAGCAG GGCTGCTAGC TTGCTTTCTG 2400  
CATTTCAAA ATGATAATTT ACAATGGAAG GACAAGAATG TCATATTCTC AAGGAAAAAA 2460  
GGTATATCAT ATGTCTCAAT CTCTCAAAT ATTCCATTG CAGACAGACC GTCATATTCT 2520  
AATAGCTCAT GAAATTGGG CAGCAGGAG GAAAGTCCCC GAAATTAATA AAATTAAAA 2580  
CTCTTATGTC AAGATGTTGA TTTGAAGCTG TTATAAGAT TGGGATTCCA GATTTGTAAA 2640  
AAGACCCCA ATGATTCTGG ACACATAGAT TTTTGTGGG GAGGTTGGC TTGAACATAA 2700  
ATGAATATC CTGTATTTC TTAGGATATC TGGTTAGTA AATTATAATA GTAGAAATAA 2760  
TACATGAATC CCAATCACAG GTTCTCAGC CCAAGCAACA AGSTAATTGC GTGCCATCA 2820  
GCATCTCAC AGAGCAGACA ACCTATTGA GGAACCAAC TGAATCCAC CTCTCTTC 2880  
ACACTGAGCC CTCTCTGATT CCTCCGTGTT GTGATGTGAT GCTGGCCAGC TTCCCAAACG 2940  
GCAGCTCCAC TGGGTCCCCT TGGTGTGTAG GACAGGAAAT GAAACATTAG GAGCTCTGCT 3000  
TGGAAACAG TCCACTACT AGGGATTTT GTTCTCTAAA ACTTTTATT TGAGGAGCAG 3060  
TAGTTTTCTA TGTTTAATG ACAGAACTG GCTAATGGA TTACAGAGG TGTTCAGCG 3120  
TATCACTGTT ATGATCTGT GTTTAGATTA TCCACTCATG CTCTCCTAT TGTACTGCAG 3180  
GTGTACTTGA AAAGTGTCC CAGTGTACTT GAACAGTTGC ATTTATAAG GGGGAAATGT 3240  
GGTTAATGG TGCTGTATAT CTCAAAGTCT TTTGTACATA ACATATATAT ATATATACAT 3300  
ATATATAAAT ATAAATATA ATATATCTCA TTGCAGCCAG TGATTAGAT TTACAGCTTA 3360  
CTCTGGGGTT ATCTCTCTGT CTAGAGCATT GTTGTCTTTC ACTGCAGTCC AGTGGGATT 3420  
ATTCCAAAAG TTTTGTAGT CTTGAGCTG GCGTGTGGCC CCGCTGTGAT CATACCCTGA 3480  
GCACGACGAA GCAACCTCGT TTCTGAGGAA GAAGCTTGAG TTCTGACTCA CTGAAATGCG 3540  
TGTGTGGTTG AAGATATCTT TTTTCTTTT CTGCTCACCC CTTTGTCTC CAACCTCCAT 3600  
TTCTGTTCAC TTTGTGGAGA GGGCATTACT TGTTCGTTAT AGACATGGAC GTTAAGAGAT 3660  
ATTCAAACT CAGAAGCATC AGCAATGTTT CTCTTTTCTT AGTTCAATCT GCAGAAATGGA 3720  
AACCCTATGC TATTAGAAAT GACAGTACTT ATTAATTGAG TCCCTAAGGA ATATTAGCC 3780  
CACTACATAG ATAGCTTTTT TTTTTTTTT TTTTITTTAA TAAGGACACC TCTTTCCAAA 3840  
CAGGCCATCA AATATGTTCT TATCTCAGAC TTACGTTGTT TTAAGAGTTT GGAAGATAC 3900  
ACATCTTTTC ATACCCCCC TTAGGAGGTT GGGCTTTCAT ATCACTCAG CCAACTGTGG 3960  
CTCTTAATTT ATTGCAATAT GATATCCACA TCAGCCAACT GTGGCTCTTT AATTTATTGC 4020  
ATAATGATAT TCACATCCCC TCAGTTGCAG TGAATGTGA GCAAAAGATC TTGAAAGCAA 4080  
AAGACACTAA TTAGTTTAAA ATGTCACTTT TTTGGTTTTT ATTATACAAA AACCATGAAG 4140  
TACTTTTTTT ATTGCTAAA TCAGATTGTT CCTTTTAGT GACTCATGTT TATGAAGAGA 4200  
GTTGAGTTTA ACATCTTAG CTTTTAAAAG AAACATTTTA ATGTAATAA TTCTACATGT 4260  
CATTCAGATA TTAATGATAT CTCTAGCCT TTAATCTGTA CTTTAAATGT ACATATTCT 4320  
GTCTTGGCGT ATTTGTATAT TTCAGTGGTT TAAAAACAA ACATGGAAG GCTTATTCCA 4380  
AATGGAAGAT AGAATATAAA ATAAACGTT ACTTGTAAAA AAAAAAAA

Seq ID NO: 249 Protein sequence:  
Protein Accession #: NP\_003383

85  
1 11 21 31 41 51  
| | | | | |

MAGSAMSSKP FLVALAIFPQ FAQVVEIANS WWSLQMNPNV QMSEVYIIGA QPLCSQLAGL 60  
 SQGQKLCHEL YQDHMQYIGE QAKTGIKECQ YQFRHRRWNC STVDNITSVFG RVMQIGSRET 120  
 APTXAVSAG VVNAMSACR EGELSTCGCS RAARPKDLPR DNLWGGCGDN IDYGYRFAKE 180  
 FVDARETERI HAKGSYSESAR ILMNLEHNEA GRRTVYNLAD VACKCHGVSG SCSLKTCLWQ 240  
 LADPFKVGDA LKEKYDSAAA MRLNSRGLV QVNSRPNSTP TQDLVYIDPS PDYCVRNEST 300  
 GSLGTQGRCL NRTSBSMDGC ELMCCGRGYD QFKTVQTERC HCKFWCCVY KCKKCTEIVD 360  
 QFVCK

Seq ID NO: 250 DNA sequence  
 Nucleic Acid Accession #: NM\_014058  
 Coding sequence: 56..1324

1 11 21 31 41 51  
 TGACTTGGAT GTAGACCTCG ACCTTCACAG GACTCTTCAT TGCTGGTGG CAATGATGTA 60  
 TCGGCCAGAT GTGGTGAGGG CTAGGAAAAG AGTTTGTGG GAACCTGGG TTATCGGCGCT 120  
 CGTCATCTTC ATATCCCTGA TTGTCCTGGC AGTGTGCATG GGACTCACTG TTCATTATGT 180  
 GAGATATAAT CAAAAGAAGA CCTACAATTA CTATAGCACA TTGTCATTTA CAACTGACAA 240  
 ACTATATGCT GAGTTTGGCA GAGAGGCTTC TAACAATTTT ACAGAAATGA GCCAGAGACT 300  
 TGAATCAATG GTGAAAAATG CATTTTATAA ATCTCCATTA AGGGAAGAAT TTGTCAAGTC 360  
 TCAGGTATC AAGTTCAGTC AACAGAAGCA TGGAGTGTG GCTCATATGC TGTGTATTG 420  
 TAGATTTTAC TCTACTGAGG ATCCTGAAAC TGTAGATAAA ATTGTTCAAC TTGTTTACA 480  
 TGAAGAGCTG CAAGATGCTG TAGGACCCCC TAAAGTAGAT CCTCACTCAG TTAATAATTAA 540  
 AAAAATCAAC AAGACAGAAA CAGACAGCTA TCTAAACCAT TGCTGCGGAA CACGAAGAAG 600  
 TAAAACCTTA GGTGAGAGTC TCAGGATCGT TGGTGGGACA GAAGTAGAAG AGGGTGAATG 660  
 GCCCTGGCAG GCTGCTGCTG AGTGGGATGG GAGTCATCGC TGTGGAGCAA CCTTAATTAA 720  
 TGCCACATGG CTTGTGAGTG CTGCTCACTG TTTTACAACA TATAAGAACC CTGCCAGATG 780  
 GACTGCTTCC TTTGGAGTAA CAATAAAACC TCGAAAAATG AAACGGGCTC TCCGAGAAAT 840  
 AATTGTCCAT GAAAAATACA AACACCCATC ACATGACTAT GATATTTCTC TTGCAGAGCT 900  
 TTCTAGCCCT GTTCCCTACA CAAATGCACT ACATAGAGTT TGTCTCCCTG ATGCATCCTA 960  
 TGAGTTTCAA CAGGTGATG TGAATTTTGT GACAGGATT GGAGCACTGA AAAATGATGG 1020  
 TTACAGTCAA AATCATCTTC GACAGCACA GGTGACTCTC ATAGACGCTA CAACTTGCAA 1080  
 TGAACCTCAA GCTTACAATG ACGCCATAAC TCCTAGAATG TTATGTGCTG GCTCCTTAGA 1140  
 AGGAAAAACA GATGCTGCTC AGGGTGACTC TGGAGGACCA CTGGTGTAGT CAGATGCTAG 1200  
 AGATATCTGG TACCTGCTG GAATAGTGAG CTGGGGAGAT GAATGTGCGA AACCCAAACA 1260  
 GCCTGGGTGT TATACTAGAG ITACGGCCTT GCGGGAGCTG ATTACTTCAA AAACCTGGAT 1320  
 CTAAGAGAGA AAGCCCTCAT GGAACAGATA ACATTTTCTT TTGTTTCTTG GGTGTGGAGG 1380  
 CCATTTTCTG AGATACAGAA TTGGAGAAGA CTTGCAAAAC AGCTAGATTG GACTGATCTC 1440  
 AATAAACTGT TTGCTTGATG CAAAAAATA A

Seq ID NO: 251 Protein sequence:  
 Protein Accession #: NP\_054777

1 11 21 31 41 51  
 MYRPDVVRAR KRVCWEPWVI GLVIFISLIV LAVCIGLTVH VRYNQKRTY NYYSTLSFTT 60  
 DRLVAEPGRE ASNFTFMSQ RLESMVKNFV YKSPLEEFV KQVIXFSPQ KHGVLAEHLL 120  
 ICRPHSTEDP STVDKIVQLV LHEKLQDAVG PPKVDPHSVK IKKINKTETD SYLNHCCGTR 180  
 RSKTLGQSLR IVGGEVEEAG EWPWQASLQW DGSHRCCGATL INATWLVSAA HCPFTYKNPA 240  
 RHTASPGVTI KPSRMKRLR RIIVREKYKH PSHDYDISLA ELSSFPVPTN AVHVRCLPDA 300  
 SYEFQPGDVM PVTGFGALKV DGYSQNEHLRQ AQVTLIDATT CNEPQAYNDA ITPRMLCAGS 360  
 LEGKTADACQ DSGGPLVSSD ARDIWYLAGI VSWGDECAKE NKPGVYTRVT ALRDWITSKT 420  
 GI

Seq ID NO: 252 DNA sequence  
 Nucleic Acid Accession #: NM\_003504.2  
 Coding sequence: 71-1771

1 11 21 31 41 51  
 GGCACGAGGC CTCGTGCCGC CGGGCTCTTG GTACCTCAGC GCGAGCGCCA GCGCTCCGGC 60  
 CGCGTGGCT ATGTTCTGTG CCGATTTCCG CAAAGAGTTC TACGAGGTGG TCCAGAGCCA 120  
 GAGGTCCTT CTCTCTGTGG CCTCGACGCT GGTGCTCTG TGTGCGTGA AGATCCTTCA 180  
 GGCCTTGTTC CAGTGTGACC ACGTGCATA TACGCTGGTT CCACTTCTG GGTGGCAAGA 240  
 ACTTGAAACT GCATTTCTTG AGCATAAAGA ACAGTTTCAT TATTTTATTC TCATAAACTG 300  
 TGGAGCTAAT GTAGACCTAT TGGATATTCT TCAACCTGAT GAAGACACTA TATTTCTTTG 360  
 GTGTGACACC CATAGGCCAG TCAATGTCGT CAATGTATAC AACGATACCC AGATCAAATT 420  
 ACTCATTAAG CAAGATGATG ACCTTGAAGT TCCCGCCTAT GAAGACATCT TCAGGGATGA 480  
 AGAGGAGGAT GAAGAGCAAT CAGGAAATGA CAGTGTATGG TCAGAGCCTT CTGAGAAGCG 540  
 CACACGGTTA GAAGAGGAGA TAGTGGAGCA AACCATGCGG AGGAGGCAGC GGCAGAGATG 600  
 GGAGGCCCGG AGAAGAGACA TCCTCTTTGA CTACGAGCAG TATGAATATC ATGGGACATC 660  
 GTCAGCCATG GTGATGTTTG AGCTGGCTTG GATGCTGTCC AAGGACCTGA ATGACATGCT 720  
 GTGGTGGGCC ATCGTTGGAC TAACAGACCA GTGGGTGCAA GACAAGATCA CTCAAATGAA 780  
 ATACGTGACT GATGTTGGTG TCCTGCAGCG CCACTTTTCC CGCCACAACC ACCGGAACGA 840  
 GGATGAGGAG AACACACTCT CCGTGGACTG CACAGCGATC TCCTTTGAGT ATGACCTCCG 900  
 CCTGTGCTC TACCAGCACT GGTCCCTCCA TGACAGCCTG TGCAACACCA GCTATACCCG 960  
 AGCCAGGTTT AAGCTGTGGT CTGTGCTATG ACAGAAAGCG CTCAGGAGT TCCTTGCGAG 1020  
 CATGGGTCTT CCGCTGAAGC AGGTGAAGCA GAAGTTCAG GCCATGAGCA TCTCCTTGAA 1080  
 GGAGAAATTG CGGGAATGA TTGAAGATC TGCAAAATAA TTTGGGATGA AGGACATGCG 1140  
 CGTGCAACTT TTAGCATTC ATTITGGGT CAAGCACAAG TTTCTGGCCA GCGACGTGGT 1200  
 CTTTGCCACC ATGCTTTTGA TGGAGAGCCC CGAGAAGGAT GGTCAAGGGA CAGATCACTT 1260  
 CATCCAGGCT CTGGACAGCC TCTCCAGGAG TAACCTGGAC AAGCTGTACC ATGGCCTGGA 1320  
 ACTGCCCAAG AAGCAGCTGC GAGCCACCCA GCAGACCACT GCCAGCTGCC TTTGACCCAA 1380  
 CCTGCTCATC TCCAGGGGCT CTTTCTGTA CTGCTCTCTC ATGGAGGGCA CTCAGATGT 1440  
 CATGCTGTTT TCTAGGCGCG CATCCCTAAG CTGCTCAGC AAACACCTGC TCAAGTCTT 1500  
 TGTGTGTTCC ACAAAGAACC GCGCTGCAA ACTGCTGCCC CTGGTGATG CTGCCCCCT 1560

GAGCATGGAG CATGGCACAG TGACCGTGGT GGGCATCCCC CCAGAGACCG ACAGCTCGGA 1620  
CAGGAAGAAC TTTTGTGGGA GGGCGTTTGA GAAGGCAGCG GAAAGCACCA GCTCCCGGAT 1680  
GCTGCACAAAC CATTTTGACC TCTCAGTAAT TGAGCTGAAA GCTGAGGATC GGAGCAAGTT 1740  
TCTGGACGCA CTATTATCCC TCCTGTCCCTA GGAATTTGAT TCTTCCAGAA TGACCTTCTT 1800  
ATTATGTATA CTGGCTTTCA TTTAGATTGT AAGTTATGGA CATGATTTGA GATGTAGAAG 1860  
CCATTTTTTA TTAATAAAAA TGCTTATTTT AGGCTCCGTC CCCAAAAAAA AAAAAAAA 1920  
AAAAAAA AA

Seq ID NO: 253 Protein sequence:  
Protein Accession #: NP\_003495.1

1 11 21 31 41 51  
MFVSDFRKEF YEVVQSRVL LFVASDVAL CACKILQALF QCDHVQYTLV PVSQWQELT 60  
AFLEHKEQFH YFILNCGAN VDLDLILQPD EDTIFFVCDT HRPVNVVNVY NDTQIKLLIK 120  
QDDLEVPAY EDIFRDEED EEHSGNDSOG SEPSEKRTL EEEIVBQTM RRRQREWEAR 180  
RRDILFDYEQ YEHGTSSAM VMFELAWMLS KDLNDMLNWA IVGLTDQWVQ DKITQMKYVT 240  
DVGVLQRHVS RHNHRNEDEE NTLSVDCTRI SFEDLRLVL YQHWLSLDSL CNTSYTAARF 300  
KLWVHGQKR LQEPFLADML PLKQVKQKFP AMDISLKENL REMIEESANK FGMKDMRVQT 360  
FSHFGFKHK FLASDVVAT MSLMESPEKD GSGTDHFIQA LDSLRSRENLD KLVHGLFLAK 420  
KQLRATQTI ASCLCTNLVI SQGPFLYCSL MEGTPDVMLF SRPASLSLLS KHLKSFVCS 480  
TKNRCKLLP LVMAAPLSME HGTVTVVGIP PETDSSDRKN FFGRAFEKAA ESTSSRMLHN 540  
HFDLSVIELK AEDRSKFLDA LISLLS

Seq ID NO: 254 DNA sequence  
Nucleic Acid Accession #: NM\_022337  
Coding sequence: 48..683

1 11 21 31 41 51  
GGCTGCGCTT CCTGGTCAG GCAAGGCAAG TCTGGCCGGC GCGCAGGATG CAGGCCCGGC 60  
ACAAGGAGCA CCTGTACAAG TTGCTGGTGA TTGGCGACCT GGGCGTGGGG AAGACCAGTA 120  
TCATCAAGCG CTACGTGCAC CAGAACTTCT CCTGCGACTA CCGGCCACCA ATCGGCGTGG 180  
ACTTGGCGCT CAGGTGCTC CACTGGGACC CGGAGACTGT GGTGGCGCTG CAGCTCTGGG 240  
ATATCGCAGG TCAAGAAAGA TTTGGAACA TGACGAGGGT CTATTACOGA GAAGCTATGG 300  
GTGCATTAT TGTCTCGAT GTCAACAGGC CAGCCACATT TGAAGCAGTG GCAAGTGGGA 360  
AAAAATGATT GGACTCCAAG TTAAGTCTCC CTAATGGCAA ACCGTTTCA GTGGTTTGT 420  
TGGCCACAA ATGTGACCAG GGAAGGATG TGCTCATGAA CAATGCGCTC AAGATGGACC 480  
AGTTCTGCAA GGAGCACGGT TTCGTAGGAT GGTTTGAAAC ATCAGCAAG GAAATATATA 540  
ACATTGATGA AGCCTCCAGA TGCGTGGTGA AACACATACT TGCAAATGAG TGTGACCTAA 600  
TGGAGTCTAT TGAGCCGGAC GTGCTGAAGC CCCATCTCAC ATCAACCAAG GTTGCCAGCT 660  
GCTCTGGCTG TGCCAAATCC TAGTAGGCAC CTTTGTGGT GTCTGGTAGG AATGACCTCA 720  
TTGTCCACA AATTGTGCTT CTATTTTAC CATTTTGGGT AAACGTGAGG ATAGATATAC 780  
CACATGTGCG AAGCCAAAGA TCTATGCTTC TGTTTTTCAT ATGAGAGAGA AATAGCAAT 840  
GTTCTTTCTA TGTCTTCTC ACCATCATCA CAGTGTTCAT AAACCTTTGA AATATTTAG 900  
TCTGTTACAA ACTTCTGTCA TGTAGCTGAC CAAATCTCTG CAGGGCCACA GTCCGCACTG 960  
TTAATTGCTT CTTTAAATCA GCAAGGCGCT CAAAGTCTTA AATAAAGGG GAGAAGAACA 1020  
AACTAGCTGT CAGTCAAGG ACTGGCTTTC ACCTTGCCCT GGTGTCTTT TCCAGATTTC 1080  
AATATATCT CTGATGCGCT GACAGGCTTA TTAAGTAGAT GTGATATTT CTTCAGAGT 1140  
GACCTCCATT CTCGGCAGAC CTAAGAGTTG CCTCTGAGTT AGCTCTTTGG AATCGTGAAC 1200  
ACAGGTGTGC TATATTGTCC TTGCTCAAC TGTCACTGTC CATGGCTGTA ATGTTGGCTT 1260  
AACTGAATAT TGTATGAAAA GACATGCTTC CATATGTGCC TTCTGTTAG CTCTCTTTGA 1320  
CTCAAGCTGT GGGGCTCTC TATACATGCT ATACATGTAA TATATATTAT ATATATTTT 1380  
GCAAGTGAAC AATAAACAT TAAAGATAA AA

Seq ID NO: 255 Protein sequence:  
Protein Accession #: NP\_071732

1 11 21 31 41 51  
MQAPHKEHLY KLLVIGDLGV GKTSIIKRYV HQNFSSHYRA TIGVDFALKV LHWDPETVVR 60  
LQLWDIAGQE RFGNMTRVYV REAMGAFIVF DVTRPATFEA VAKWKNDLDS KLSLPNGKPV 120  
SVVLLANKCD QGKDVLMNNG LKMDQFCKEH GFWGWFETSA KENINIDEAS RCLVKHILAN 180  
ECDLMESIEF DVVKPHLTET KVASCSCGCAK S

Seq ID NO: 256 DNA sequence  
Nucleic Acid Accession #: NM\_016321  
Coding sequence: 25..1464

1 11 21 31 41 51  
GGAACGCCCC GCTGCCAGCC CGGCCAGGCA CCCCTGCAGC ATGGCCTGGA ACACCAACCT 60  
CCGCTGGCGG CTGCCGCTCA CCTGCCTGCT CCTGCAGGTG ATTATGGTGA TTCTCTTCGG 120  
GGTGTTCGTG GCTACGACT TCGAGGCCGA CGCCCACTGG TGTTCAGAGA GGACGCACAA 180  
GAACCTGAGC GACATGGAGA ACGAATTCTA CTATCGCTAC CCAAGCTTCC AGGACGTGCA 240  
CGTGATGGTC TTGCTGGGCT TCGGCTTCTC CATGACTTTC CTGACGCGCT ACGGCTTCAG 300  
CGCCGTGGCC TTCAACTTCC TGTGGCAGC CTTCGGCATC CAGTGGGCGC TGTCTATGCA 360  
GGGCTGGTTC CACTTCTTAC AAGACCGCTA CATCGTCTGT GGCCTGGAGA ACCTCATCAA 420  
CGCTGACTTC TCGTGGGCT CTGCTGCGT GGCCTTGGG GCAGTCTTGG GTAAAGTCAG 480  
CCCATTCAG CGCTCATCA TGACTTTCTT CCAAGTGACC CTCTCGCTG TGAATGAGTT 540  
CATTCTCCTT AACCTGCTAA AGGTGAAGGA TGCAGGAGGC TCATGACCA TCACACATT 600  
TGGCGCTAC TTTGGGCTCA CAGTGACCG GATCCTTAC CGACGAAC TAGAGCAGAG 660  
CAAGGAGAGA CAGAACTCTG TGTACAGTC GGACCTCTTT GCCATGATTG GCACCTCTT 720  
CCTGTGGATG TACTGGCCCA GCTTCAACTC AGCCATATCC TACCATGGGG ACAGCCAGCA 780  
CCGAGCGGCC ATCAACACCT ACTGCTCTT GGCAGCGCTG GTGCTTACCT CGGTGGCAAT 840

ATCCAGTGGC CTGCACAGA AGGGCAAGCT GGACATGGTG CACATCCAGA ATGCCACGCT 900  
 CGCAGGAGGG GTGGCCGTGG GTACCGCTGC TGAGATGATG CTCATGCCCTT ACGGTGCCCT 960  
 CATCATCGCG TTGGTCTGCG GCATCATCTC CACCCTGGGT TTTGTATACC TGACCCCATTT 1020  
 5 CTGGAGTCC CGGCTGCACA TCCAGGACAC ATGTGGCATT AACAACTGCG ATGGCATTCC 1080  
 TGGCATCATA GGGCGCATCG TGGGTGCTGT GACAGCGGCC TCCGCCAGCC TTGAAGTCTA 1140  
 TGGAAAGAA GGGCTTGTCC ATTCCTTTGA CTTTCAAGGT TTCAACGGGG ACTGGACCCG 1200  
 AAGAACACAG GGAAAGTTCC AGATTATGG TCTCTGGTG ACCCTGGCCA TGGCCCTGAT 1260  
 GGGTGGCATC ATTTGGGGC TCATTITGAG ATTACCATTC TGGGGACAAC CTTCAGATGA 1320  
 10 GAATCTGCTT GAGGATGCGG TCTACTGGGA GATGCTGAA GGGACACAGA CTGTCTACAT 1380  
 CCTGAGGAC CCCACCTTCA AGCCCTCAGG ACCCTCAGTA CCCTCAGTAC CCATGGTGTG 1440  
 CCCACTACCC ATGGCTTCTT CGGTACCCCTT GGTACCCCTAG GCTCCAGGG CAGGTGAGGA 1500  
 GCAGGCTCCA CAGACTSTCC TGGGGCCAG AGGAGCTGGT GCTGACCTAG CTAGGGATGC 1560  
 AAGAGTGAGC AAGCAGCACC CCCACCTGCT GGCTTGGCCT CAAGGTGCTT CCACCCCTGC 1620  
 15 CCTCCCTTC ATCCAGGGG GTCTGMCCTGA GAATGGAGAA GGGAAGCTA CAAAGTGGGC 1680  
 ATCCAAGCG GTTCTGGCT GCAGAAGTTC TGCTCTGCC TGGGGTCTTG GCCACATTGG 1740  
 AGAAAAACAG GCTCAAGTGC GGGCTGGGAC CTGGTGGGTG AACCTGAGCT CTCCAGGAG 1800  
 ACAACTTAGC TGCCAGTAC CACCTATGAG GCTCTTCTAC CCGGTGCTG CACCTCGGCC 1860  
 AGCATCTCTT ATGCTCCCTG GGTCCCCAG ACCTCTCTGT GTTGTGTGCG TGGCAGCCTC 1920  
 20 CAGGAATAAA CATTCTTGT GTCTTTGTA AAAAAAAAA AAAAAAAAA

Seq ID NO: 257 Protein sequence:  
 Protein Accession #: NP\_057405

1 11 21 31 41 51  
 MANNTNLRNR LPLTCLLLQV IMVILPGVFP RYDFEADAHW WSETRHKNLS DMENEFYYRY 60  
 PSFQDVHVMV FVGSGFLMTF LQRYGFSAVG FNFLLAAGFI QWALLMQGWF HFLQDRYIVV 120  
 GVENLINADF CVASVCVAPG AVLKGVSPHQ LLIMTFQVIT LFAVNEFILL NLLKVKDAGG 180  
 30 SMTIHTFGAY FGLTVTRILY RRLNLEQSKER QNSVYQSDLE AMIGTLFLWM YWPSFNSAIS 240  
 YHGDQSHRAA INTYCSLAAC VLTSVAISSA LHKKGKLDLV HIQNTATLGG VAVGTAAEMM 300  
 LMPYGLIILG FVYLITPFLS RLEIQDTGCI NNLHGIPGII GGIVGAVTAA 360  
 SASLEVYKKE GLVHSFDFQG FNGDWTARTQ GKFIYGLLV TLMALMGGI IVGLILRLPF 420  
 WGQPSDENCF EDAVWEMPE GNSTVYIPED PTFKPSGFSV PSVPMVSPLE MASSVPLVP

Seq ID NO: 258 DNA sequence  
 Nucleic Acid Accession #: NM\_002358.2  
 Coding sequence: 75..692

1 11 21 31 41 51  
 GGGAAAGTGT GTTGGAGCCG CTGTGGTTGC TGTCCGCGGA GTGGAAGCGC GTGCTTTTGT 60  
 TTGTGTCCCT GGCCATGGCG CTGCAGCTCT CCGGGAGCA GGGAAATCACC CTGCGCGGGA 120  
 45 GCGCGAAAT CGTGGCCGAG TTCTTCTCAT TCGGCATCAA CAGCATTTTA TATCAGCGTG 180  
 GCATATATCC ATCTGAAACC TTACTCGAG TGCAGAAATA CGGACTCACC TTGCTTGTA 240  
 CTACTGATCT TGAGCTCATA AAATACCTAA ATAAATGGGT GGAACAACTG AAAGATTGGT 300  
 TATACAAGTG TTCACTTCAG AAATCTGGTG TAGTTATCTC AAATATTGAA AGTGGTGAGG 360  
 TCCCTGGAAG ATGGCAGTTT GATATTGAGT GTGACAAGAC TGCAAAAGAT GACAGTGCAC 420  
 50 CCAGAGAAAA GTCTCAGAAA GCTATCCAGG ATGAAATCCG TTCAGTGATC AGACAGATCA 480  
 CAGCTACGGT GACATTCTCG CCACTGTGCG AAGTTTCTTG TTCATTGAT CTGCTGATTT 540  
 ATACAGACAA AGATTGGTTT GTACCTGAAA AATGGGAAGA GTGCGGACCA CAGTTTATTA 600  
 CCAATTCTGA GGAAGTCCCG CTTCGTTTCT TACTACTTAC AATCCACAAA GTAATATGCA 660  
 TGGTGGCCTA CAAAATTCCT GTCAATGACT GAGGATGACA TGAGGAAAAA AATGTAATTG 720  
 55 TAATTTTGA AATGTGGTTT CTGAAATCA GGTATCTAT AGTTGATATG TTTTATTTCA 780  
 TTGGTTAATT TTTACATGGA GAAACCAAA ATGATACTTA CTGAACGTG TGTAAATTGT 840  
 CCTTTATTTT TTTGTACCTT ATTTGACTTA CCATGGAGTT AACATCATGA ATTTATTGCA 900  
 CATTTGTTCA AAGGAACCGG GAGGTTTTTT TGTAACATT GTGATGTATA TTCTTTTGAA 960  
 60 GATAGTAACT TAGATGGGAA AAACCTGTGC TATAAGCTA GATGCTTTCC TAAATCAGAT 1020  
 GTTTTGGTCA AGTAGTTTGA CTCACTATAG GTAGGGAGAT ATTAAAGTAT AAAATACAAC 1080  
 AAAGGAAGTC TAAATATICA GAATCTTTGT TAAGGTCTCG AAAGTAACCT ATAATCTATA 1140  
 AACCAATGAA TATTGCTGTA TAGCTCCTTT TGACCTTCAT TTCATGTATA GTTTTCCCTA 1200  
 TTGAATCAGT TTCCAATTAT TTGACTTTAA TTTATGTAAC TTGAACCTAT GAAGCAATGG 1260  
 65 ATATTGTGAC TGTTTAATGT TCTGTGATAC AGAACTCTTA AAAATGTTTT TTCATGTGTT 1320  
 TTATAAAATC AAGTTTAAAG TGAAAGTGAG GAAATAAAGT TAAAGTTTGT TTAATAAAAA 1380  
 AAAAAAAAA

Seq ID NO: 259 Protein sequence:  
 Protein Accession #: NP\_002349.1

1 11 21 31 41 51  
 MALQLSREQG ITRGSABIV ABFFSFGINS ILYQRGIYPS ETPTRVQXYG LITLLVTTDL 60  
 75 LIKYLNVVE QKLDWLYKCS VQKLWVVISN IESGEVLERW QFDIECDKTA KDDSPAPREKS 120  
 QKAIQDEIRS VIRQITATVT FLPLLEVSCS FDLIIYTDKD LVVPEKWEES GPQFITNSEE 180  
 VRLRSFTTTI HKVNSMVAYK IPVND

Seq ID NO: 260 DNA sequence  
 Nucleic Acid Accession #: NM\_001211  
 Coding sequence: 43..3195

1 11 21 31 41 51  
 AAAGGCCTGC AGCAGGACGA GGACCTGAGC CAGGAATGCA GGATGGCGGC GGTGAAGAG 60  
 85 GAAGGGGGTG CTCTGAGTGA AGCCATGTCC CTGGAGGGAG ATGAATGGGA ACTGAGTAAA 120  
 GAAAATGTAC AACCTTTAAG GCAAGGGCGG ATCATGTCCA CGCTTCAGGG AGCACTGGCA 180  
 CAAGAATCTG CCTGTAAACA TACTCTTCAG CAGCAGAAAC GGGCATTTGA ATATGAAATT 240

CGATTTTACA CTGGAAATGA CCTCTGGAT GTTTGGGATA GGTATATCAG CTGGACAGAG 300  
 CAGAACTATC CTCAAGGTGG GAAAGAGAGT AATATGTCAA CGTTATTAGA AAGAGCTGTA 360  
 GAAGCACTAC AAGGAGAAAA AGCATATTAT AGTGATCCTC GATTTCTCAA TCTCTGGCTT 420  
 AAATTAGGGC GTTTATGCAA TGAGCCTTGG GATATGTACA GTTACTTGCA CAACCAAGGG 480  
 ATTGGTGTIT CACTTGCTCA GTTCTATATC TCATGGGCAG AAGATATGA AGCTAGAGAA 540  
 AACTTTAGGA AAGCAGATGC GATATTTTCG GAAGGGATTG AACAGAAGGC TGAACCACTA 600  
 GAAAGACTAC AGTCCCGACA CCGACAATTC CAAGCTCGAG TGTCTCGGCA AACTCTGTTG 660  
 GCACCTTGAGA AAGAAGAAGA GGAGGAAGTT TTTGAGTCTT CTGTACCACA ACGAAGCACA 720  
 CTAGCTGAAC TAAAGAGCAA AGGGAAAAAG ACAGCAAGAG CTCCAATCAT CGGTGTAGGA 780  
 GGTGCTCTCA AGGCTCCAAG CCAGAACAGA GGACTCCAAA ATCCATTTC TCAACAGATG 840  
 CAAAAATAA GTAGAAATAC TGTTTTGAT GAAAATGCTG ATGAGGCTTC TACAGCAGAG 900  
 TTGTCTAAGC CTACAGTCCA GCCATGGATA GCACCCCCCA TGCCCGAGGC CAAAGAGAAT 960  
 GAGCTGCAAG CAGGCCCTTG GAACACAGGC AGGTCTCTGG AACACAGGCC TCGTGGCAAT 1020  
 ACAGCTTCAC TGAAGCTGT ACCCGCTGTG CTTCCAGTT TCACTCCATA TGTGGAAGAG 1080  
 ACTGCACAA ACAGGATTTT GACACCATGT AAAATTGAAC CTAGTATAAA CCACATCCTA 1140  
 AGCACAGAA AGCCTGGAAA GGAAGAAGGA GATCCTCTAC AAAGGGTTCA GAGCCATCAG 1200  
 CAAGCGTCTG AGGAGAAGAA AGAGAAGATG ATGTATTGTA AGGAGAAGAT TTATGCAGGA 1260  
 GTAGGGGAAT TCTCCTTTGA AGAAATTCGG GCTGAAGTTT TCCGGAAGAA ATTAAAGAG 1320  
 CAAAGGGAAG CCGAGCTATT GACCACTGCA GAGAAGAGAG CAGAAATGCA GAAACAGATT 1380  
 GAAGAGATGG AAGAAGAGCT AAAAGAAATC CAACTACTCT AGCAAGAAAG AACAGGTGAT 1440  
 CAGCAAGAA AGACCATGCC TACAAAGGAG ACAACTAAAC TGCAAAATTC TCCGAGTCT 1500  
 CAGAAATAC CAGGAATGAC TCTATCCAGT TCTGTTTGTG AAGTAAACTG TTGTGCCAGA 1560  
 GAACTTCAC TTGCGAGAAA CATTTGGCAG GAACAACCTC ATTCTAAAGG TCCAGTGTGA 1620  
 CCTTTCTCCA GTTTTGTATG GTTCTCTCTT TCAGAAAGA AGAATAAAG TCTCTGTGA 1680  
 GATCCCCCAG GAGTTTATG TCAACGAAGA CCCCTTGCAG TTCTCAAAC CTCGAAAGC 1740  
 ATCACCCTCA ATGAAGATGT GTCTCCAGAT GTTGTGATG AATTACAGG AATTGAACCC 1800  
 TTGAGCGAGG ATGCCATTAT CACAGGCTTC AGAAATGTAA CAATTGTGCC TAACCCAGAA 1860  
 GACACTTGTG ACTTTGCCAG AGCAGCTCGT TTTGTATCCA CTCCTTTTCA TGAGATAATG 1920  
 TCTTTGAAGG ATCTCCCTTC TGATCCTGAG AGACTGTTAG CGGAAGAAGA TCTAGATGTA 1980  
 AAGACCTCTG AGGACCAAGC GACAGCTTGT GGCATATCT ACAGTCAGAC TCTCAGCATC 2040  
 AAGAAGCTGA GGCCTAATTAT TGAAGACAGT CGTGAAGCCA CACACTCCTC TGGCTTCTCT 2100  
 GGTCTCTCTG CTTCCGTTGC AAGCACCTCC TCCATCAAA GTCTTCAAAT TCTGAGAAA 2160  
 CTAGAACCTA CTAATGAGAC TTCAGAAAAC CCTACTCAGT CACCATGGTG TTCACAGTAT 2220  
 CGCAGACAGC TACTGAAGTC CCTACCAGAG TTAAGTGCCT CTGCAGAGTT GTGTATAGAA 2280  
 GACAGACCAA TGCCCTAAGT GGAATTTGAG AAGGAATTTG AATTAGGTAA TGAGGATTAC 2340  
 TGCATTAAAC GGCCTAATCT AATATGTGAA GATTACAGT TATTCTGGGT GCGCCCAAGA 2400  
 AACTCTGACG AATTAAAGT AATAAAGGTA TCTTCTCAAC CTGTCCCATG GGACTTTTAT 2460  
 ATCAACCTCA AGTTTAAAGGA ACGTTTAAAT GAAGATTTTG ATCATTTTTG CAGCTGTTAT 2520  
 CAATATCAAG ATGGCTGTAT TGTTTGGCAC CAATATATAA ACTGCTTAC CCTTCAGGAT 2580  
 CTTCTCCAAC ACAGTGAATA TATTACCAT GAAATAACAG TGTGATTAT TTATAACCTT 2640  
 TTGACAATAG TGAAGATGCT ACACAAAGCA GAAATAGTCC ATGGTGACTT GAGTCCAAGG 2700  
 TGTCTGATTC TCAGAAACAG AATCCAAGAT CCCTATGATT GTAACAAGAA CAATCAAGCT 2760  
 TTGAAGATAG TGAATTTTC CTACAGTGT GACCTTAGGG TGCAGCTGGA TGTTTTACC 2820  
 CTCAGCGGCT TTCGAGCTGT ACAGATCCTG GAAGGACAAA AGATCCTGGC TAACTGTCTT 2880  
 TCTCCCTACC AGGTAGACCT GTTTGGTATA CGAGATTTAG CACATTTACT ATTGTTCAAG 2940  
 GAACACCTAC AGGTCTTCTG GGATGGGTCC TTCTGGAAC ITAGCCAAA TATTCTGAG 3000  
 CTAAAGATAG GTGAATTTG GAATAAATTC TTTGTGCGGA TTCTGAATGC CAATGATGAG 3060  
 GCCACAGTGT GTTCTCTGG GGAGCTTGCA GCAGAAATGA ATGGGGTTTT TGACACTACA 3120  
 TTCCAAAGTC ACCTGAACAA AGCCTTATGG AAGGTAGGGA AGTTAACTAG TCTTGGGCT 3180  
 TTGCTCTTTC AGTAGCTAG GCAATCAAGT CTCACAGATT GCTGCCTCAG AGCAATGGTT 3240  
 GTATTGTGGA ACAGTGAAC TGTATGTGCT GTAATTTAAT TTAGGACACA TTTAGATGCA 3300  
 CTACCATGTC TGTCTCACTT TTTGTACAG GTATATTTG ACGTCACTGA TATTTTTAT 3360  
 ACAGTATAT ACTTACTCAT GGCCTTGTCT AACTTTTGTG AAGAAGTATT TTATTCTAAA 3420  
 CAGACTCATT ACAAATGTTT ACCTTGTAT TTAACCCATT TGTCTCTACT TTTCCCTGTA 3480  
 CTTTTCCTCAT TTGTAATTTG TAAATGTTT TCTTATGATC ACCATGTATT TTGTAATAA 3540  
 TAAATAGTA TCTGTTAAAA AAAAAA AAA

Seq ID NO: 261 Protein sequence:  
 Protein Accession #: NP\_001202

1 11 21 31 41 51  
 MAAVKKBGGA LSEAMLEGD EWELSKENVQ PLRQGRIMST LQGALAQESA CNNTLQQQKR 60  
 AFPEYIRFYT GNDPLDVWDR YISWTEQNY P QGGKESNMST LLERAVEALQ GEKRYYSDDR 120  
 FLNLMLKLR LCNELDMYS YLHNGIGVS LAQFYISWAE EYEARENFRK ADAIFQEGIQ 180  
 QKAEPLERLQ SQHRQFQARV SRQTLLEAK EEEBEVFESS VPQRSTLAEK KSKGKKKTARA 240  
 PIIRVGGALK APSQNRQLQN PFPQMQNNS RITVFDENAD EASTAELSKP TVQPWIAAPP 300  
 PRAKENELQA GPWNTGRSLE HRPRGNTASL IAVPAVLPSF TPVVEETAQK PVMTPCKIEP 360  
 SINHLSTRK PKKEEGDPLQ RVQSHQQA SE EKKELMYCK EKIYAGVGEP SFEEIRAEVF 420  
 RKKLKEQREA ELLTSAEKRA EMQKQIEEME KKLKEIQTTQ QERTGDQEE TMPTKETTKL 480  
 QIASESQKIP GMTLBSVVCQ VNCCARETSL AENIQEQPH SKGPSVPFSI FDEPLLSEKK 540  
 NKSPPADPPR VLAQRRLPLV LKTSSESITSN EDVSPDVCE FTGIEPLSED AIITGFRNVT 600  
 ICPNFEDTCD FARAARFVST PFHEIMSLKD LPSDPERLLP EEDLDVKTSE DQQTACGTIY 660  
 SQTLSIKKLS PIIEDSREAT HSSGFSGSSA SVASTSSIKC LQIPEKLELT NETSENPTQS 720  
 PWCQYRRLK LKSLPELSAS AELCIEDRPM PKLEIEKEIE LGNEDYCIKR EYLICEDYKL 780  
 FNVAPRNSAE LTVIKVSSQP VPWDFYINLK LKERLNEDFD HFCSCYQYQD GCIVWHQYIN 840  
 CPTLQDLQH SEYITHEITV LIYMLLTIV EMLHKAIEVH GDLSPRCLIL RNRIHDPYDC 900  
 LKNNQALKIV DFSYVDLIV QLDVFTLSGF RTVQILEGQK ILANCSFPYQ VDLFGIADLA 960  
 HLLLFKEHLQ VPWDSFVKL SQNISELKDQ ELWNKFPVRI LNANDEATVS VLGLAEAEVN 1020  
 GVFDITTFQSH LNKALWKVKG LTSPGALLFQ

Seq ID NO: 262 DNA sequence  
 Nucleic Acid Accession #: NM\_003784  
 Coding sequence: 365..1507

1 11 21 31 41 51

	GTCTACTTAT	CAATAAGCAG	CTGCCTGTGC	AGAGTGCAGG	CTGCACCTTT	GGACAGCCTT	60
	TAAACTGAA	TTCTCAGAA	TTTAGAACAA	ATTTTGTCT	AGAAATGCTG	ACTTTGGTTC	120
5	ATTAGTAGT	GGTAAACAG	GCTCCCTTCG	AAGCTCTCCT	TCATCACCTT	CCTAAGTGCA	180
	TGTACAGGA	AGCTCTCCTT	CATCACCTTC	CTAAGTGCAT	GGGGGAAAT	ACCTAGGGCT	240
	CAACAGTCTT	GAGAAGTGTG	GAAACATTTT	CTTTGTGAGT	GAGAACAGAT	CACCTAGAGA	300
	AAGGAACCA	GATTCCCATC	ACTGCTTCTG	GGTATCAGAT	GCTAGCGCTG	CACTCCATTT	360
	TGCAATGGCC	TCCTTGTCTG	CAGCAAAATG	AGAGTTTTCG	TTCAACCTGT	TCAGAGAGAT	420
10	GGATGACAA	CAAGGAAATG	GAAATGTGTT	CTTTTCTCT	CTGAGCCTCT	TCGCTGCCCT	480
	GGCCCTGGTC	CGCTTGGGCG	CTCAAGATGA	CTCCCTCTCT	CAGATTGATA	AGTTGCTTCA	540
	TGTTAACACT	GCCTCAGGAT	ATGGAACCTC	TTCTAATAGT	CAGTCAGGGC	TCCAGTCTCA	600
	ACTGAAAAGA	GTTTTTCTG	ATATAAATGC	ATCCCAACA	GATTATGATC	TCAGCATTTG	660
	GAATGGGCTT	TTTGCTGAAA	AAGTGTATGG	CTTTCATAAG	GACTACATTG	AGTGTGCCGA	720
	AAAAATTATC	GATGCCAAAG	TGGAGCGAGT	TGACTTTACG	AATCATTTAG	AAGACACTAG	780
15	ACGTAATATT	AATAAGTGGG	TTGAAAATGA	AACACATGGC	AAAATCAAGA	ACGTGATTGG	840
	TGAAGTGGC	ATAAGCTCAT	CTGCTGTAAT	GGTGTCTGGT	AATGCTGTGT	ACTTCAAAGG	900
	CAAGTGGCAA	TCAGCCTTCA	CCAAGAGCGA	AACCATAAAT	TGCCATTTC	AATCTCCCAA	960
	GTGCTCTGG	AAGGCAGTCG	CCATGATGCA	TCAGGAACGG	AAGTTCAATT	TGCTCTGTTAT	1020
20	TGAGGACCCA	TCATGAAGA	TTCTTGAGCT	CAGATACAA	GGTGGCATA	ACATGTACGT	1080
	TCCTGCTGCT	GAGAATGACC	TCTCTGAAAT	TGAAAACAAA	CTGACCTTTC	AGAATCTAAT	1140
	GGAATGGACC	AATCAAGGC	GAATGACCTC	TAAGTATGTT	GAGGTATTTT	TTCCCTAGTT	1200
	CAAGATAGAG	AAGAATTATG	AAATGAAACA	ATATTGAGA	GCCCTAGGGC	TGAAAGATAT	1260
	CTTTGATGAA	TCCAAGCAG	ATCTCTCTGG	GATTGCTTCG	GGGGTCTGTC	TGTATATATC	1320
25	AAGGATGATG	CACAAATCTT	ACATAGAGGT	CAGTGGAGG	GGCACCGAGG	CTACTGCTGC	1380
	CACAGGAATG	AATATTGTAG	AAAAGCAACT	CCCTCAGTCC	ACGCTGTTTA	GAGCTGACCA	1440
	CCCATTCCTA	TTTGTATATC	GGAAGGATGA	CATCATCTTA	TTCACTGGCA	AAGTTTCTTG	1500
	CCCTTGAAAA	TCCAATTGGT	TTCTGTTATA	GCAGTCCCCA	CAACATCAAA	GRACCAACAC	1560
	AAGTCAATAG	ATTTGRTT	AATTGAAAA	ATGTGGTGT	TCCTTTGAGT	TTATTCTTTC	1620
30	CTAACATTGG	TCAGCAGATG	ACACTGGTGA	CTTGACCTCT	CTAGACACCC	TGGTTGATTG	1680
	TCCTGATCCC	TGCTCTTAGC	ATCTTACCAC	CATGTGCTC	ACCCATTCT	AATTTCATTG	1740
	TCCTTCTTCC	CAGCTCATTT	TCTATCATT	TCCCCATGA	CCCGCTGGA	AATTATGGAG	1800
	RGTCCTCAAC	TGGTAAGGAG	AACGTAGAAG	TAGCCCTAGG	GATCCTTTT	GAAACTCTAC	1860
	AGTTATCGCA	GATATTCTAG	CTTCATTGTA	AGCAATCTAG	GAAATAAGCC	CTGCTGCTTT	1920
35	CTAGAAATAA	GTGTGAAGGA	TAAATTTTCT	TTGTGACCT	ATGAAGATT	TAGAGTTTAC	1980
	CTTCATATGT	TTGATTTTAA	ATCAGTGTAT	AATCTAGATG	GTAATAAATG	TGAATTTGGG	2040
	ATTAGGGACC	TACCAAAATA	TTTCATTAA	GCTTTCAATT	GACAAATTTT	GGCCTTCTTT	2100
	TGATAGACA	ATATGTACAT	GTTTTTTCAA	ATATTAAAGA	TCITTTAACT	GTGGCAGTT	2160
40	GTTATCTACA	GAATCATATT	TCATATGCTG	TGTAGTTTAT	AAGTTTITCC	TCATTTATC	2220
	AGAATAAAGA	AATACAACAT	ACCTGTAAA				

Seq ID NO: 263 Protein sequence:  
Protein Accession #: NP\_003775

45	1	11	21	31	41	51	
	MASLAAANAE	PCFNLFREMD	DNQNGNVFF	SSLSLFAALA	LVRIGAQDD	LSQIDKLELV	60
	NTASGYGNSS	NSQSLQ9QL	KRVFSDINAS	HKDYDLSIVN	GLFAEKVYGF	HKDYIECAEK	120
50	LYDAKVERVD	FTMHLEDTRR	NINKNVENET	HGKIKNVIGE	GGISSAVMV	LVNAVYFKGK	180
	WQSAPTKSET	INCHFKSPKC	SGKAVAMMHQ	ERKFNLSVIE	DPSMKILELR	YNGGINMYVL	240
	LPENDLSIEI	NKLTQNLME	WTFPRMTSK	YVEVFPPQFK	IEKNYEMKQY	LRLALGLKDF	300
	DESKADLSGI	ASGRLYISR	MMHKSYLEVT	EEGTETAAT	GSNIVBKQLP	QSTLPRADHP	360
	FLFVIRKDDI	ILFSGKVSCE					

55 Seq ID NO: 264 DNA sequence  
Nucleic Acid Accession #: AB052906  
Coding sequence: 74-814

60	1	11	21	31	41	51	
	AAAACCTTGA	GGTGATTTCAT	CTTCCAGGCT	CTCCTTCCAT	CAAGTCTCTC	CTCCCTAGCG	60
	CTCTGGGTCC	TTAATGGCAG	CAGCCGCGCG	TACCAAGATC	CTTCTGTGCC	TCCCGCTTCT	120
	GCTCCTGTCT	TCCGGCTGGT	CCCGGGCTGG	GCGAGCCGAC	CCTCACTCTC	TTTGCTATGA	180
65	CATCACCGTC	ATCCCTAAGT	TCAGACCTGG	ACCAACGGTG	TGTGCGGTTT	AAGGCCAGGT	240
	GGATGAAAAG	ACTTTTCTTC	ACTATGACTG	TGGCAACAAG	ACAGTCACAC	CTGTCAGTCC	300
	CCTGGGGAAG	AACTAAATG	TCACAAACGC	CTGGAAGCA	CAGAACCCAG	TACTGAGAGA	360
	GGTGGTGGAC	ATACTTACAG	AGCAACTGCG	TGACATTGAG	CTGGAGAATT	ACACACCCAA	420
	GGAAACCCCTC	ACCTTGCAGG	CCAGGATGTC	TTGTGAGCAG	AAAGCTGAAG	GACACAGCAG	480
70	TGGATCTTGG	CAGTTTCAGT	TCGATGGGCA	GATCTTCTCT	CTCTTTGACT	CAGAGAAGAG	540
	AATGTGGACA	ACGGTTTCAT	CTGGAGCCAG	AAAGATGAAA	GAAAAGTGGG	AGAATGACAA	600
	GGTTGTGGCC	ATGTCCTTCC	ATTACTTCTC	AATGGGAGAC	TGTATAGGAT	GGCTTGAGGA	660
	CTTCTTGATG	GGCATGGACA	GCACCTTGGA	GCCAAGTGCA	GGAGCACCAC	TGCGCATGTC	720
	CTCAGGCACA	ACCCAACCTC	GGGCCACAGC	CACCAACCTC	ATCCTTTGCT	GCCTCTCAT	780
75	CATCCTCCCC	TGCTTCATCC	TCCCTGGCAT	CTGAGGAGAG	TCCTTTAGAG	TGACAGGTTA	840
	AAGCTGATAC	CAAAAGGCTC	CTGTGAGCAC	GGTCTTGATC	AAACTCGCCC	TTCTGTCTGG	900
	CCAGCTGCCC	ACGACCTACG	GTGTATGTCC	AGTGGCCTCC	AGCAGATCAT	GATGACATCA	960
	TGGACCCAA	AGCTCATTC	CTGCCTTGAT	TCCTTTTGCC	AACAATTTTA	CCAGCAGTTA	1020
	TACCTAACAT	ATTATGCAAT	TTTCTCTTGG	TGCTACCTGA	TGGAATTCCT	GCACTTAAAG	1080
80	TTCTGGCTGA	CTAAACAAGA	TATATCATTT	TCCTTTCTCT	CTTTTGTGTT	GGAAAATCAA	1140
	GTACTTCTTT	GAAATGATAT	CTCTTTCTTG	CAAATGATAT	TGTCAGTAAA	ATAATCACGT	1200
	TAGACTTCAG	ACCTCTGGGG	ATCTCTTCCG	TGTCCTGAAA	GAGAATTTT	AAATTTATTA	1260
	ATAAGAAAA	ATTTATATTA	ATGATTGTTT	CCTTTAGTAA	TTTATTGTTT	TGTACTGATA	1320
	TTTAATAAAA	GAGTTCTATT	TCCCAAAAA	AAAAAATAA			

85 Seq ID NO: 265 Protein sequence:  
Protein Accession #: BAB61048.1

1 11 21 31 41 51  
MAAAATKIL LCLPLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGGQVDEKT 60  
FLHYDCGNKT VTPVSPGLGKK LNVTTAMKAQ NFVLRREVVDI LTELRLDIQL ENYTPKEPLT 120  
LQARMSCEQK ABGHSSGSWQ PSFDGQIFLL FDSKRMWTT VHPGARJOMKE KWENDKVVM 180  
SPHYFSMGDC IGWLEDFLMG MDSTLEPSAG AFLAMSSGTT QLRATATTLI LCCLLIILPC 240  
FILPGI

Seq ID NO: 266 DNA sequence  
Nucleic Acid Accession #: XM\_084853.1  
Coding sequence: 127-444

1 11 21 31 41 51  
ATTGATGATA TATTAAACGA AATCAAATTT GGTGAATATG TGGACACTGG AAAGCTAATC 60  
GACAAGATCA ACTTACCAGA TTCTCTAAAA GTGTACCTTA ACCACAAGCC ACCTTTTGGT 120  
AACACCATGA GTGGCATCCA CAAGAGCTTT GAGGTGCTCG GTTATACCAA CTCCAAAGGG 180  
AAAAAGGCCA TTCGAAGAGA GGACTTCCTG AGACTGCTCG TTACTAAAGG TGAGCATATG 240  
ACGGAGGAGG AGATGTTGGA TTGCTTTGCT TCACGTGTTG GCCTGAATCC CGAGGGATGG 300  
AAATCCGAGC CTGCAACCTG CTCGCTCAAA GGTTCAGAAA TTGCTCTTGA AGAAGAACTT 360  
CCAGACGAAA TCACGTGAGA AATATTCGCG ACTGAATTC TTGCTTAAC CATTTCAGAA 420  
GATTCGGGCC AGGATGGTCA GTGAAGTTAC CAGGAATGTT TAAAGCACA AGGACTTTGG 480  
GTGTGTGTGC ATGCACATGT GTGTGTTTTC CATGAGGCAC TGCTTTTAT GCATTTCCCT 540  
CCCCCTCTC ATCTTAGAA CATTAGACA TTAAAGCAAG TTTCTGGTGA GCAATG

Seq ID NO: 267 Protein sequence:  
Protein Accession #: XP\_084853.1

1 11 21 31 41 51  
MSGIHKSEFEV LGYNSKGGK AIRREDFRLR LVTKGEHMT EEMLDGFASL FGLNPEGWKS 60  
EPATCSVKGS EICLEELPD EITAEIFATE ILGLTISEDS GQDQG

Seq ID NO: 268 DNA sequence  
Nucleic Acid Accession #: NM\_001898  
Coding sequence: 57-482

1 11 21 31 41 51  
GGCTCTCACC CTCCTCTCCT GCAGCTCCAG CTTTGTGCTC TGCTCTGAG GAGACCATGG 60  
CCCAGTATCT GAGTACCCTG CTGCTCCTGC TGGCCACCT AGCTGTGGCC CTGGCCTGGA 120  
GCCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCAATGATG 180  
AGTGGGTACA GCGTGCCCTT CACTTCGCCA TCAGCGAGTA TAACAGGCC ACCAAGATG 240  
ACTACTACAG ACGTCCGCTG CGGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGGTGA 300  
ATTACTTCTT CGAGTAGAG GTGGGCGCA CCATATGTAC CAAGTCCAG CCCAAGCTGG 360  
ACACCTGTGC CTTCATGAA CAGCCAGAAC TGCAAGAGA ACAGTGTGC TCTTCGAGA 420  
TCTACGAAGT TCCTGGGAG AACAGAAGGT CCTGTGTGAA ATCCAGGTG CAAGAATCCT 480  
AGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCCAC TCCACCCCC TGTAGTGTCT 540  
CCACCCCTGG ACTGTGGCC CCAACCTGC GGGAGGCTCT CCATGTGCC TCGCCCAAGA 600  
GACAGACAGA GAAGGTGCA GAGTCCCTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660  
CTTCTCTCTT GCTTCTAATA GCCCTGTAC ATGGTACACA CCCCCCACC TCCTGCAATT 720  
AAACAGTAGC ATCGCC

Seq ID NO: 269 Protein sequence:  
Protein Accession #: NP\_001889.1

1 11 21 31 41 51  
MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQALHF AISEYNKATK 60  
DDYYRRLPRV LRARQQTVGG VNYFFDVEVG RTICTKSQPN LDTCAFHEQP ELQKKQLCSF 120  
EIVEVPWENR RSLVKSRCQE S

Seq ID NO: 270 DNA sequence  
Nucleic Acid Accession #: XM\_093210  
Coding sequence: 13-1854

1 11 21 31 41 51  
ATGGCAAGCG CGGAATCTC CTCAGCTGCC GTTTCACAAA AGAGGTACCA GGTCCGCACC 60  
AAACGAGCAC ACAAGCAGA CCAGGAGCTG CAGAAGAAGG AGGCGGCAGC GATGGACCAG 120  
GGCAGAGGGA ATGGGAGGG GGCATCTTAC CCCATATCTG AGGTGCGACT GCGGAGCTA 180  
GAGCGGACTG GGCCTTTCCC GTTGGCGGT GGCCTCAATC AGGACTTCTT GCCCAGGTGC 240  
GCCTTCAAAA CGGTAAGAGC TGCAACTGAA CGTGTGAGAC ATGGTGCAGA TAGGCTGAGA 300  
GGCGGGGGGA GAGATGCCCA TGAAGTCAAG TACCGGACA CGCCCTCCAC TTCTACCACC 360  
ACGAGTAACA CCGCCCCAC GGGACCGCTC TCGAGGTCCC CCAAGCCAA GACGCAAGGA 420  
GGAAACCCCC GCGCGCGCGC CAGCAGCGGC GGGCACCGGC CCAATGGCCA CGGAAGTCA 480  
CACTGGCAGT CCGCCCTCCT CACACCGCAG GCGTGCACTG TGGCGAGCG AGCCTCCCG 540  
GCGAGGAGCC AGCTAGGCC GTCAACCGG TTGCTCCAC GGAAGGGGC ACCAGGCAAA 600  
CTGCCAAGG CCGCAGGCC AGGCTCCCTG GCGGAGGCT CCGTGTGTC CCGCCAGATC 660  
ATGGCGGCA CAGAGCTCCC GAGCCATGGC TTCTGTCCG GGAAGGGCC GCGCTCTCG 720  
CTGTCCAGCT AG

Seq ID NO: 271 Protein sequence:  
Protein Accession #: XP\_093210

1 11 21 31 41 51



MLRHGEQKRR RARKKNDPLF TCAFKTVRAA TERVRHGADR LRGGGRDAHE LKYPDTPSTS 60  
TTTSNTAPTG PLRSRSPKPRP QGGTPRRRPA AAGTRANGHG TQHWQSALLT PQACSVADGA 120  
5 SRAPDPARPS PRLLPREGAG GKLPKAPSPG SLAEASAGLL AHVRLQNAQA QRVISIQALP 180  
PNSSVGRKEE RPAGQQRRA PAFMATLST GSRPSSHRRR AVWPTEPPGP RTQLEPSRL 240  
LPREGAPGKL PKAPSPGSLA EASAGPAQIM AATRLPSRGP LSGNGPASWL SS

Seq ID NO: 272 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..732

1 11 21 31 41 51  
GGATACTGTG TCACTCAAAG TAATGGGAGG GAGAGAGAAC AGGAGGGGTA GGGATGCTTT 60  
15 TGAAAAAGCT TTTTTCCTCA CTTTTAACCT GCTTTAGCGT TAAGAGTACT TACCAGCTAA 120  
TAATGTGGAG GAAATTATTC TTCTCATTTG GAGATTACAG AATATATCTA TTCATCTTGA 180  
ATACCCACTT GAAGCCTCTG TAGAAATGTC TCGTCTCCG GTTGTATTTT TAAACCTTAC 240  
ATGATTTTGT CTGTGTTCTG CAGTGAGAAA TTACATCCAT AGCAAAGACA AAAGTCTTTT 300  
20 TAAATTATTT TTATTATCTT TTCATATAGT TCTTACAATT TCTAAAAAT TAACACTCAT 360  
TTAGTATCAC AATTATAGGG AGAGGGTTTT TTGTATTTTT AAGCATATGT GGCTTATATA 420  
AAAAATTGAG AAGTCATAGG ACTGTCAATG ATTGCAGCTC TGAGAACCAA TGCCCTGAAAC 480  
TTAAGCC

Seq ID NO: 273 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
MGGRENREGR DAFKAPFPPT FNLL

Seq ID NO: 274 DNA sequence  
Nucleic Acid Accession #: NM\_003976.2  
Coding sequence: 299-961

1 11 21 31 41 51  
CTCTGAGCTT CTCTGAGCCT TGTITGCTCA TCTGGAAAAA GGGGATTAAA CCATTACCT 60  
40 CATGGAGTTG TGAAAGAATA GCTGCAAGC ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120  
CTACTTCTGC TGGGTGAGT CTAGCTGTGT AGGCCCTGTG TTCCTCACCT GGAGAACTG 180  
GGGTGGCAGG CCGGTCCCCC ACAAAGATA ACTCATCTCT TAATTGCAA GCTGCCCTCA 240  
CAGGAGGTGA GGGGAACAGC TCAACAATGG CTGATGGGCG CTCCTGGTGT TGATAGAGAT 300  
GGAACCTTGA CTGTGAGGCC TCTCCACGCT GTCCCACTGC CCTGCGCTTA GCGGCGAGCC 360  
45 TGCCCTGTGG CCCACCTGG CCGCTCTGGC TCTGCTGAGC AGCGTCCGAG AGGCTCCCT 420  
GGCTCCCGG CCGCCAGCC CTGCCCCCG CGAAGGCCCG CCGCTGTGTC TGGGTCCCC 480  
CGCCGCCAC CTGCGGGGG GACGCACGGC CCGCTGGTGC AGTGAAGAG CCGCGCGGCC 540  
GCCGCCGAG CCTTCTCGGC CCGGCCCCCC GCGCGCTGCA CCCCCTCTG CTCTTCCCCG 600  
CGGGGGCGG GGGGCGCGG CTGGGGGGCC GGGCAGCCGC GCTGGGCGAG CCGGGGCGG 660  
GGCTGCGCG CTGCGCTCGC AGCTGGTGCC GGTGCGCGCG CTGCGGCTGG GCCACCGCTC 720  
50 CGACGAGCTG GTGGGTTCG GCTTCTGCG CCGCTCCTGC CGCGCGCGCG GCTCTCCACA 780  
CGACCTCAGC CTGCGCAGCC TACTGGGCGC CCGGGCCCTG CGACCGCCCC CCGGCTCCCG 840  
GCCGTCAGC CAGCCCTGCT GCCGACCCAC GCGCTACGAA GCGGTCTCCT TCATGGACGT 900  
CAACAGCAC CAGGCAACCT TGGACCGCT CTCCGCCACC GCGTGGGCTG GCTTGGGCTG 960  
AGGGCTCGCT CAGGGCTTT GCAGACTGGA CCTTACCGG TGGCTCTTCC TGCCCTGGAC 1020  
55 CTTCCGCTG GTGCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCTT CAAAGCTGAG 1080  
AGGCCCTAC CCGTGGGTGA TGGATATCAT CCGCGAACAG GTGAAGGAC AACTGACTAG 1140  
CAGCCCCAGA GCCCTCACCC TGGGATCCC AGCCTAAAG ACACCAGAGA CCTCAGCTAT 1200  
GGAGCCCTT GAGCCCTTT CTCACAGACT CTGGCACTGG CCAGGCGCTG AACCTGGGAC 1260  
60 CCTCCTCTG ATGAACACTA CAGTGGCTGA GGCATCAGCC CCGGCCAGG CCTGTAGGG 1320  
ACAGCATTTG AAGGACACT ATTGCAGTTG CTTGGTTGAA AGTGCCCTGT CTGGAACCTG 1380  
CCTGTACTCA CTCATGGAG CTGGCCCC

Seq ID NO: 275 Protein sequence:  
Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
70 PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
RGCRRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPFGS 180  
RPVSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 276 DNA sequence  
Nucleic Acid Accession #: NM\_057091.1  
Coding sequence: 783-1445

1 11 21 31 41 51  
ACTGGCCGCT GAGAGAAGAA TCGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60  
80 GGACCCCAAA ATCTGCACGT ACCAGCAGTC AGCCGCCCCA CGCAGGGACC GCTTACCCC 120  
TGCTTCCCG CCTTCACTCA CTCTCTCCG CCTCGGCCG GGCTTCCAG CTCTCTACTT 180  
CGGTGTCTA CAACTCAAC TCCCGTTTC CGTGCCTCTC CACGCTCGA GTTCTCTACT 240  
CTCCTATCC GAGGGGCCCC TCCAGCATC TACCCCTCTC CCACTCTCG GGGACCTAGC 300  
85 CAGCTAGGG GGGACTGGAT CCGAGGGGTG GAGCAGCCAG GTGAGCCCCG AAAGGTGGGG 360  
CGGGGAGGG GGGTCCCGG CCGCAGCCG GGTCTGGTG AGCTGGGGG TGGAATTGGA 420  
CACCGGAGG CTGCGGCGG GGGCAGGAG CTGCTGAGG ATGGAGTTG GCCCGCCCC 480  
CAGACAAGG CCGGGGCTC CGCCAGCAGC AGGTCCCTCG GGGCCAGCC CTCGCTGCCA 540

CCCGGGCGCTG GAGCCCCACA CCCGAGGGTG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600  
 TAAAAGAGGC ACTGCCAGGT GTACAGTCCT GGGCATGGCG TGTTTGAGCT TCGGGGGAGA 660  
 GCCCAGCACT GGTCCCCGGA AAGGTGCCCTA GAAGAACAAG GTGCAGGACC CCGTGTCTGC 720  
 TCAACAGGAG GGTGGGGGAA CAGCTCAACA ATGGCTGATG GGCGCTCCTG GTGTGTATAG 780  
 AGATGGAACT TGGACTTGA GGCCTCTCCA CGCTGTCCCA CTGCCCTCGG CCTAGGCGGC 840  
 AGCCTGCCCT GTGGCCACC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCCT 900  
 CCTGGGCTC CGCGCCCGC AGCCTTGCC CCGCGAAGG CCCCCCGCT GTCTGGCGT 960  
 CCCCAGCGG CCACTGCGG GGGGAGCGCA CGGCCCGCTG GTGCAGTGA AGAGCCCGGC 1020  
 GGCGCGCGC CGAGCTTCT CGGCCCGGC CCCCAGCGC TGACCCCCA TCTGCTCTTC 1080  
 CCGCGCGGG CGCGCGCGC CGGGCTGGG GCGCGGCGC CGCGCTCGG GCAGCGGGG 1140  
 CGCGGGGCTG CGCCTGCGC TCGCAGCTGG TGCCGCTGG CGCGCTCGG CTGGGCCACC 1200  
 GCTCCGACGA GCTGGTGGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCG GCCTGCTCTC 1260  
 CACACGACTC CAGCTGCGC AGCCTACTGG GCGCGGCGC CCGCGACCG CCCCAGGCT 1320  
 CCGCGCGCT CAGCCAGCCC TGCTGCCGAC CCACGCGCTA CGAAGCGGTC TCCTTCATGG 1380  
 AGCTCAACAG CACTTGAGA ACCGTGGACC GCCTCTCCGC CACCGCTCG GGCTGCTTGG 1440  
 CTGAGGGCT CGCTCCAGG CTTTGACAG TGGACCCCTA CCGTGGGCTC TTCTGCTCTC 1500  
 GGACCTCCG GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGACGAAG GCCTCAAGC 1560  
 TGAGAGGCCC CTACCGTGG GTGATGGATA TCATCCCGA ACAGGTGAAG GGCAACTGA 1620  
 CTAGCAGCCC CAGAGCCCTC ACCCTGCGGA TCCAGCCTA AAAGACCA GAGACCTCAG 1680  
 CTATGGAGCC CTTGGAGCC ACTTCTACA GACTCTGGCA CTGGCCAGGC CTGCAACTTG 1740  
 GGACCTCTC CTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGC CAGGCCCTGT 1800  
 AGGGACAGCA TTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGGC TGTGCTGGAA 1860  
 CTGGCCTGTA CTACTCATG GGAGCTGGCC CC

Seq ID NO: 277 Protein sequence:

Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
 MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
 PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPALP RGGRAARAGG PGSRRARAAGA 120  
 RGRRLRSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPPGS 180  
 RFSVQPCRP TRYBAVFMVD VNSTRWTVDR LSATACGCLG

Seq ID NO: 278 DNA sequence

Nucleic Acid Accession #: NM\_057160.1

Coding sequence: 1-714

1 11 21 31 41 51  
 ATGCCCGGCC TGATCTCAGC CGAGGACAG CCCCTCCTTG AGGTCTCTCC TCCCCAAGCC 60  
 CACTGGGTG CCTCTTTCT CACTTGCTC TCTCCGCGCA GCCTGCCCTG 120  
 TGGCCACCC TGGCGCTCT GGCCTCTGCT AGCAGGCTCG CAGAGGCCTC CCTGGGCTCC 180  
 GCGCCCGCA CGCTTGCCC CCGCGAAGGC CCCCAGCTG TCTGGCGTC CCCCAGCGC 240  
 CACTGCGCG GGGAGCGCAC GCGCGCTGG TGCACTGGA GAGCCCGCG GCGCGCGCG 300  
 CAGCCTTCTC GCGCGCGCC CCGCGCGCT GCACCCCAT CTGCTCTTCC CCGCGGGGG 360  
 CGCGCGCGC GGGCTGGGG CCGCGGCGC CGCGCTCGG CAGCGGGGC GCGGGGCTG 420  
 CGCTTGCGT CCGAGCTGGT GCGGCTGCG GCGCTCGGC TGGGCCACG CTCCGACGAG 480  
 CTGCTGCGT TCGCTTCTG CAGCGGCTCC TGCCGCGCG CCGCTCTCC ACACGACCTC 540  
 AGCCTGGCCA GCTACTGGG GCGCGGGGC CTGCGAGCG CCCCAGGCTC CCGGCCGCTC 600  
 AGCCAGCCCT GCTGCGGACC CACGCGCTAC GAAGCGTCT CCTTCATGA CGTCAACAGC 660  
 ACCTGGAGAA CCGTGGACCG CCTCTCGCC ACCGCTGGG GCTGCTTGG CTGAGGGCTC 720  
 GCTCCAGGGC TTGCGAGCT GGACCCCTAC CGTGGGCTC TCTGCTTGG GACCTTCCG 780  
 CAGAGTCCCA CTAGCCAGC GCCTCAGCCA GGGACGAAG CCTCAAGCT GAGAGGCCCC 840  
 TACCGGTGGG TGATGGATAT CATCCCGAA CAGGTGAAG GACAACCTG TAGCAGCCCC 900  
 AGAGCCCTCA CCGTGGGAT CCGAGCCTAA AAGACACAG AGACCTCAG TATGGAGCCC 960  
 TTGGGACCCA CTCTCAGAG ACTCTGGAC TGGCCAGGC TCGAAGCTG GACCCCTCT 1020  
 CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGCC AGGCCCTGTA GGGACAGCAT 1080  
 TGAAGGACA CATATTGAG TTGCTTGGT GAAAGTGCT GTGCTGGAAC TGGCCTGTAC 1140  
 TCATCTATG GAGCTGCCCC C

Seq ID NO: 279 Protein sequence:

Protein Accession #: NP\_476501.1

1 11 21 31 41 51  
 MPGLISARG PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SVAEASLGS 60  
 APRSPAPREG PPPVLASPA HLPGRRTARW CSGRARRPPF QPSRPAPPPP APPSALPRGG 120  
 RAARAGPGS RARAAGARG RLRSQLVVPR ALGLGHRSD LVRFRFCGSG CRRARSPHDL 180  
 SLASLLGAGA LRPPPGSRPV SQPCRPTRY EAVSFMDVNS TWRTVDRISA TACGLG

Seq ID NO: 280 DNA sequence

Nucleic Acid Accession #: NM\_057090.1

Coding sequence: 29-715

1 11 21 31 41 51  
 CTGATGGCG CTCTGGTGT TGATAGAGAT GGAAGTGA CTGGAGGCC TCTCCAGCT 60  
 GTCCACTGC CCTGGCCTA GGGGAGGAG TCACTTGT CTCTCGCGC AGCCTGCCCT 120  
 GTGGCCACG CTGGCGCTC TGGCTCTGCT GAGCAGGCT GCAGAGGCT CCTTGGGCTC 180  
 CGCGCCCGC AGCCTGCCC CCGCGAAGG CCCCAGCTG GTCTGGGCT CCCCAGCGC 240  
 CCACCTGCG GGGGAGCGA CCGCGCGCTG GTGCAGTGA AGAGCCCGC GCGCGCGCG 300  
 GAGCCTTCT CCGCGCGCC CCGCGCGCC TGACCCCCA TCTGCTCTC CCGCGGGGG 360  
 CCGCGCGCG CCGGCTGGG GCGCGGCGC CCGCTCGG GCAGCGGGG CCGCGGGCTG 420  
 CGCCTGCGC TCGAGCTGG TGCGGCTGG CCGCTCGG CTGGGCTC GCTCCGAGA 480  
 GCTGGTGGT TTCCGCTTCT GCAGCGGCTC CTGCGCGCG GCGCGCTCT CACAGACCT 540  
 CAGCCTGGC AGCCTACTG GCGCGGGGC CCGCGAGCG CCCCAGGCT CCGCGCGCT 600  
 CAGCCAGCCC TGCTGCGAC CCAAGCGCTA CGAAGCGGTC TCCTTCATG ACGTCAACAG 660

WO 02/086443

PCT/US02/12476

CACCTGGAGA ACOGTGGACC GCGTCTCGCG CACCGCCTGC GGCTGCCTGG GCTGAGGGCT 720  
 CGCTCCAGAG CTTTGCAGAC TGGACCCCTA CCGGTGGCTC TTCTGCTG GGCACCTCCC 780  
 GCAGAGTCCC ACTAGCCAGC GGCTCAGCC AGGGACGAAG GCCTCAAGC TGAGAGGCC 840  
 CTACCGGTGG GTGATGGATA TCATCCCGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900  
 CAGAGCCCTC ACCCTGCGGA TCCAGCCCTA AAAGACACCA GAGACCTCAG CTATGGAGCC 960  
 CTTGGGACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTCGAACCTG GGCACCTCCC 1020  
 TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCCCGCC CAGGCCCTGT AGGGACAGCA 1080  
 TTTGAAGGAC ACATATTGCA GTTGTGGT TGAAGTGCC TGTGCTGAA CTGGCCTGTA 1140  
 CTCACCTCATG GGAGCTGGCC CC

Seq ID NO: 281 Protein sequence:  
Protein Accession #: NP\_476431.1

1 11 21 31 41 51  
 MELGLGLST LSHCPWPRRQ AFLGLSAQPA LNPFLAALAL LSSVAEASLG SAPRSPAPRE 60  
 GPPFVLASPA GHLPGRRTAR WCGRRARRPP PQSPRPAPP PAPPSPALPRG GRAARAGGPG 120  
 SRARAGARG CLRLSQLVPV RALGLGHRSD ELVRFRCPSG SCRRARSPHD LSLASLGLAG 180  
 ALRPPPSRP VSQPCRPRTR YEAVSPMDVN STWRTVDRLS ATAQCLG

Seq ID NO: 282 DNA sequence  
Nucleic Acid Accession #: Eos sequence

1 11 21 31 41 51  
 CTACTGCACC TGCCCTCTGT TTCTTTTGA AATCTCTTAC CTTTCATTAG GGTTCCTTC 60  
 ATAGCAATTT CCTTTGGTTT TTAAGACTTC TACATTGCTT TTCTTTTAT TATCTGTGCT 120  
 COGTGAACCT TATGAATGCT GCTTAAAAAT AATGTCAAAA TATGTTTTAG CTGCCCTACT 180  
 AGTAAAGTT TTCTTTTGT CTATCTTGG TTTCCATATA CTATTTTGG TTTTGTGTA 240  
 GATCTAATCA ATGATCTAGT CAGAAGCTAC TTCACTGSC TACAGTGATC ATGTTTCTGT 300  
 GCTAAAAATG AACTTGAAC ACGGAAGTAG TGGTTGGTCC AGTTTGAAG CTCTTATTAG 360  
 TATTTCTCAT CCTGGCTGTA ATAATAGCCA TTATTTGTTA TGCCCTTTGT ATGTAGCAGA 420  
 CACTCTTAAG GATTTTATGT GTATTATTCA AATTGCTATT ACTGTTCTTT TTATAGTTGA 480  
 GAATCTCAGG ATACCTACAT TTATCACTTT TTCAATATAT ATGTATTCTT TATT

Seq ID NO: 283 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 564-1481

1 11 21 31 41 51  
 GAGACTTTTA ATCATCTATC CCTTGTGCTT TACGACAGCC CTACAATACA CTAGAGGCTT 60  
 CAAAGAGGTC AAAAATTCAC ATGTGTAGAC AAATTAGGTC CCTTAAGATG CCAGGCAAAC 120  
 GAAGTGCTAC CAAACACGCG AATGACTGTC CTAAGAGTGC GTTCTGGGAT ACACCTGTAA 180  
 ACTTGGATCA AGTTCCTCTC CCTCTCTCA AAATATATCG ACTTGTGCTG AAAGAAATCA 240  
 CGACCGATGC TCACAAATCT GACCTCGTAA TTATATAGGG GGTGGTTTGT GTTCTGCGT 300  
 CTTTCCCTGA TTCAGTGCCA GGTAAACATAT TTCAATGACA AAATGAACCT CAACACCACG 360  
 GCACAAACAG CAGAGGCCCT CAAAGTTGTC GGTAGGGAGC CAGGACCCCG CCAGTGGGCT 420  
 GGGGAGACAC CGTACTAAAC AAGCTTGCAC ACAGCAGGCA CCTTCTGTCC ACTGAGGAGG 480  
 AAGGGCTGGC TAAGGGAGGC CGGGGCGGAG GAAGCCAAAG TCTGCAGGCC CTGACAAAGT 540  
 CCTCCCGGCC TCCACGCGTC GCCATGGCAA CGCGGGGTCT GTGCTGGCCG GGATTTGGCG 600  
 GCCTGGCGCG CGCAGGGGCC GCTGGGAAAG CGCTCCCGCG CGCGGGCTCC GCCAGTTTGA 660  
 ACTTGGCGGG CCAGATGTGG GCGGGCGGGC GCTGGGGGCC TACTTTTCCC TCTTCTTACG 720  
 CGGGTTTCTC TGCTGACTGC AGACCCAGGT CTCGGCCCTC CTCGGACTCC TGCTCAGTCC 780  
 CTATGACGGG CGCACGTGGG CAGGGGCTGG AGGTGGTGGC CTCGCCCTCG CGCGCGCTGC 840  
 CGCTGAGCTG CAGCAATTC ACCAGGTGCG TGTGTCTCC CCTTGGCCAC CAGAGCTTCC 900  
 AGTTTGACGA GGACGACGCT GACGGGAGG ATGAGGAAGA CGTGGATGAT GAGGAAGACG 960  
 TGGATGAAGA TGCCCATGAT TCAGAGGCCA AAGTGGCGAG CCTGAGAGGA ATGGAGTTAC 1020  
 AGGGGTGCGC CAGCACTCAG GTTGAATCAG AAAATAACCA AGAAGAACAG AAACAGGTGC 1080  
 GCTTACACGA AAGCGGCTG ACACCATGGG AGGTGTGGTT TATTGGCAAA GAAAAAGAA 1140  
 AACGTGACCG GCTGCAACTG AAAGCTCTAG AGGAATTAAA TCAACACTA GAAAAAGAA 1200  
 AAGAAATGSA AGAAGCTGAA AAAAGAAAGA TAATTGCTGA AGAAAAACAC AAGGAATGGG 1260  
 TTCAGAAAAA GAATGAGCAA AAAAGAAAG AAAGAGACCA AAAAATTAAT AAAGAAATGG 1320  
 AGAAAAAGC AGCAAGGAA CTGGAGAAAG AATACTTGCA AGAAAAAGCA AAAGAAAAAT 1380  
 ATCAAGAAATG GTTAAAGAAA AAAAATGCTG AAGAATGTGA GAGGAAGAAAG AAAGAAAAAG 1440  
 AAAACAACAG CAAGCTGAAA TACAGGAGAA AAAGGAAATA GCAGAAAAAA AGTTTCAAGA 1500  
 ATGGTTGGAA AATGCGAAAC ATAAACCTCG TCCAGCTGCA AAGAGCTATG GTTATGCCAA 1560  
 TGGAAAACTT ACAGGTTTTT ACAGTGGAAA TTCTATCCA GAACACGCTT TTTATAATCC 1620  
 AATTCCTGCG AAACCAATTC ATATGCCACC TCCCAAGAA GCTAAGGATC TATCAGGAAG 1680  
 GAAGAGTAAA AGACTGTGA TAAGTCAGCC ACACAGTCA TCATCTCTGG TAATTCATAA 1740  
 AGCCAGGAGC AATCTTTGCC TTGGAACTCT GTGCAGATA CAAAGATAGC GTATGTGAAA 1800  
 AATAACATGC TTTTATCTGG AGCTATTAA TTTAAAAATC AGAAATTTGT TTTTACTGCT 1860  
 CAGTCAATAA CTCACACTT AATGTGATTA TTGACAAATA GCAATTTTGT CATTTGTATA 1920  
 TGGAGTCTCT AGAGTTGAGG AAGATATTTT CTGGATTTTG GTTTTATATA ACTTTTAAAG 1980  
 GTTGATCTTG GCATGTTGTT TTGACAAATA AGTGGCTGAA TATGTAAGAA TTGTGTTTGT 2040  
 ATTAGCTTG TATTAAGAT ACACGTGAAT ACCAATAAAA CTAACAATT TTCTTG

Seq ID NO: 284 Protein sequence:  
Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MATRGLCNFG LAGLARAGPA GKARPRRGS SLNLAGQMW AGRWGPTFPS SYAGFSADCR 60  
 PRSRPSSDSC SVPMTGARGQ GLEVVRSPP PLPLSCSNST RSLSPPLGHQ SFQFDEDDGD 120  
 GEDEEDVDDE EDVDEADHS EAKVASLRGM ELQGCSTQV ESENNQBEQK QVRLPESRLT 180  
 PWEVNFQKE KEERDRQLQK ALEELNQLE KRKEMEEREX RKI IAEKHKH EWVQKKNQEK 240  
 RKERBQKINK EMEEKAAKEL EKEYLQEKAK EKYQEWLKKK NAECECKKK EKKNSKLKY 300

RRKRX

Seq ID NO: 285 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1-1746

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

1 11 21 31 41 51  
| | | | |  
ATGCACTGA AGCATTATCT CTTTGTCTG GTGGGCTGCC AAGCCTGGGG TGCAGGTTG 60  
GCCTACCATG GCTGCCCTAG CGAGTGTACC TGCTCCAGGG CCTCCCAGGT GGAGTGCACC 120  
GGGGCAGCA TGTGGCGGT GCCACCCCT CTGCCCTGGA ACGCCATGAG CCTGCAGATC 180  
CTCAACACGC ACATCACTGA ACTCAATGAG TCCCGTTCC TCAATATCTC AGCCCTCATC 240  
GCCTGAGGA TTGAGAAGAA TGAGCTGTCT CGCATCACGC CTGGGGCCTT CCGAAACCTG 300  
GGCTCGCTGC GCTATCTCAG CCTCGCCAA AACAAGCTGC AGGTTCTGCC CATCGGCTCT 360  
TTCCAGGGCC TGACAGAGCT TGAGTCTCTC CTCTGTCCA GTAACAGCT GTTGACAGATC 420  
CAGCGGGCCC ACTTCTCCCA GTGCAGCAAC CTCAGGAGC TGCAAGTGA CCGCAACCAC 480  
CTGGAATACA TCCTGACGG AGCCTTCGAC CACCTGGTAG GACTCACGAA GCTCAATCTG 540  
GGCAAGAATA GCCTCACCA CATCTCACCC AGGGTCTTCC AGCACCTGGG CAATCTCCAG 600  
GTCTCCGGC TGTATGAGAA CAGGCTCAG GATATCCCA TGGGCATTT TGAATGGCTT 660  
GTTAACTGC AGGAACCTGC TCTACAGCAG AACAGATTG GACTGTCTC CCTGTGTCTC 720  
TCCACAACA CCGCAACCTC CCAGAGACTC TACCTGTCCA ACAACCACAT CCCCAGCTG 780  
CCACCAGCA TCTTCATGCA GTGCCCCAG CTCACCGTC TTAATCTCTT TGGGAATTCC 840  
CTGAAGGAGC TCTCTCTGGG GATCTTCGGG CCCATGCCCA ACCTGCGGGA GCTTTGGCTC 900  
TATGACAACC ACATCTCTTC TCTACCGGAC AATGTCTTCA GCAACCTCCG CCAGTTGCAG 960  
GTCTCTGATC TTAGCCGCAA TCAGATCAGC TTCATCTCCC CGGGTGCTT CAACGGGCTA 1020  
ACGGAGCTTC GGGAGCTGTC CCTCCACACC AAGCACTGC AGGACCTGGA CGGGAATGTC 1080  
TTCGGATGT TGGCAACCTC GCAGAACATC TCCTGCGACA ACAATCGCTC CAGACAGCTC 1140  
CCAGGGAATA TCTTCGCCAA GTCATGGG CTTATGGCCA TCCAGCTGCA GAACAACAG 1200  
CTGGAGAACT TGCCCTCGG CATCTTCGAT CACCTGGGGA AACTGTGTGA GCTGCGGCTG 1260  
TATGACAATC CGTGAGGTTG TGACTCAGAC ATCCTTCCCG TCCGCAACTG GCTCTGTCTC 1320  
AACCAGCTTA GGTTAGGAGG GGACACTGTA CCTGTGTGTT TCAGCCGAGC CAATGTCCGA 1380  
GGCCAGTCCC TCATTATCAT CAATGTCAAC GTTGTGTGTC CAAGCGTCCA TGTCCCTGAG 1440  
GTGCTAGTT ACCCAGAAAC ACCATGGTAC CCAGACACAC CCAATTAACC TGACACCACA 1500  
TCCGTCTCTT CTACCACTGA GCTAACACAG CCTGTGGAAG ACTACACTGA TCTGACTACC 1560  
ATTGAGGCTA CTGATGACGC CAGCGTTTGG GGCATGACCC AGGCCAGAG CGGGCTGGCC 1620  
ATTGCGGCA TTGTAATTGG CATTTGTGCT CCTGGCTGCT CCTGGCTGCT CTGCTGGCTG 1680  
TGTGTCTGCT GCAAGAGAG GAGCCAGCT GTCTGTATGC AGATGAAGGC ACCCAATGAG 1740  
TGTAAAGAG CAGGGCTGGA GCAGGGCTGG GGAATGATGG GACTGGAGGA CTTGGGAATT 1800  
TCATCTTCTT GCCTCCACCC CTGGGTCCAT GGAGCTTTCC CGTGATTGCT CTTTCTGGCC 1860  
CTAGATAAAG GTGTGCTTAC CTCTTCTGTA CTGCTCTGAT TCTCCGCTAG AGAAGCAGT 1920  
CTGTGCGGAC CTGCTCACTA TCAGGAAGAT AGATCCAAC TGGCCATGGCA AAGCCCTGG 1980  
GGATTTCGGA TTCATACCCC TGGGCTTCTC TCGAGAGGGC TCTTCTTCCA AATCCTCCCC 2040  
ACCTGTCTCT CAAGAACAGC CTTCCTTGGG CCGAGGCCCT CTCGGGGCTC CTGTAGACTC 2100  
AGTTAGTCCA CAGCCTGCTC ACTTGTGGG AATAGTCTCT CGCTGAGATA GCCCTCTCG 2160  
CCTAAGTATT ATGTAAGTTG ATTTCCCTTC TTTTGTGTTT CTGTGTTGTG CTATGCGCTG 2220  
ACCCAGCATG TCCCTCAAA TGAAAGTTCT CCGCTTGATT TTCTGCTCCT GAAGGCAGGG 2280  
TGAGTTCTCT CCTCAAAGAA GACTTCAAAC CATTAACTG GTTCTTAAAG AGCCGTCAAT 2340  
CAGCCTGGTT TTGGGATGAG TATGAAAGAG AGAAGGAAAA TCATGCGGCT CAGTTCTGCG 2400  
AGACAGAGA GCGGTATCA GTGTCTCACT TGTGATTTT ATCTGAAAAA GGAAGAAACA 2460  
CCCCAGCACA GCAAGCTCAG CCTTTTAGAG AAGGATATT CCAACTGCA AACTTTGCTT 2520  
TGAAAGTTT AGCCCTTTAA GGAATGAAAT CATGTAGAA TTTGGACTTC TAAAAACATT 2580  
AAATCAGCT TATTATACG GGATAGAGAA AGAAATCTGG TGCTGGGGG TCCCTGTGTT 2640  
CACCCTAGA GTTTGTGTTA AAATTTTAA TTGAAGCATG TGAAGTGTAC STGCAGAAAA 2700  
GTGGGAACAT GATAGTGTAT GGCTTGGTGG ATTTTCACAA ACTGAACATA CCTGTGAAT 2760  
CAGCATCTAG ACCCAGACCC AGAGCATCAC AAATATCCCC CATCTGGGC TTTTCCAGA 2820  
GGAGATGGGG GCTTCTGAAG ATGGACTTAC CTGGGACCTG CCCCCATGA GCCAGGACGG 2880  
TCCCCCACA GTGAGCTGT GCAAGGGCCC CGTGGCCAGG GGTGGAGGAG AATATGTGGG 2940  
TGTGGACAGG ATGGGAGAT GTGGCTGAA CAGGAGATT TATTATATCT GGAGACCCCTG 3000  
AGAGACCTG AGACCTGGGG CACCATGGCT GGCAGGTCG GAAGCATCTC GACTGCAGAG 3060  
GTCTGGGGC TCCCTGGAGC TTCCCTGCCA GCAAGTTGTC TGCGGCTCAT CGGAGGCCCC 3120  
TCGCTCTGGA GCCTTCTATG GACGTGATAT GCCTGTATCT GTTTTAAATT TTCATTCTTC 3180  
ACTTAGGGGA AGTGAATCG CTCAGAGATG AGATCTTTTA ATTGAAAACG AAGTGTAAAG 3240  
GAATCTAGTG TCTTTCTAAT GTGTAAAAAT TCTCCATCAA CATCACAGTC AGCTGGCAGC 3300  
TGAATCTCAG AATCTCACT ACAGCAGGCG ACACGGGGGT ACACCGATGG GTCACTACTG 3360  
GTCTGGGGC TCCCTGGAGC TCCTCTGCG TGTGTCTGG TTAGGAGTTG AGTTGTTTGC 3420  
TCCAGGGTTA TTCTCTCTCT CGAGTCACAG TCACACGAAT ACCTGCTTC TCTGGCTTTC 3480  
CTGCTATACA CATATTACA TGGCGCTCAA GAAGTTAGCG TCATGGCAAC GTGTGCTTTT 3540  
CTCTGGACAA CTGGCCAGT TTACAGTGAA ATGGAGAAIT TCAGGTCTCC ACCTCTGCCC 3600  
AGGAAAGAAC TTCAGCTGAC TCCACGGGGA TCTGGAATTC CACGACCAAT CCGATCGGC 3660  
TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGCTTTGG AAATCCACCA CCAATCCCGA 3720  
TCGGCTCTTA TTAGCTCCCC GCTCCAAG ACACCTGTGA TCTGGAATTC TACCACCAAT 3780  
CCGATCGGC TCTTATTAGC TCCCGCTCC ACAAGACACC TGTGACATCC TCCAGGGCCA 3840  
CAGGAGCAG TGCTGACCAG TTTTCCCTTC CAGTTCTCTG ACAAAAGTG TCCAGAGGGC 3900  
TGTGTGAAA CACTAGTGCA CTTTGTAGCT TTTCAACCTC TGTCCAGGG AATCTAGGAG 3960  
AGATGAGGCC CAGCAGAGTC AAGAGATGTC ATCCCCCAG GGTCTCCAAG GCATTTCCAC 4020  
ACTATTGGTG GCACCTGGAG GACATGCACC AAGGCTTGCC AGAGCCAACA GGAAGTGAGC 4080  
CCAGAGCATG GCACATGAGC ATCAACCGCT GATGGTGGCC TGCTGTGCTC GTGCGCAACA 4140  
GGGGCATCCC GGGCGTACC CCTCCAGACA GGAAGCATGG GTTTGCCAC AGACCTGTCT 4200  
GGTGTCTCTG TGAATGGCTC CCAGATGTCT TTGTGCATAG GCACAAGTGG GCCAGGGCTG 4260  
GAGGAGGTG GGAAGCTTCA TCATCGGTG GGGCTGGCA ATCTTAAACC AGAACCTTAA 4320  
GATATTCTCT GCAGTAGCCA TGACATTGGA GCACCTTCTC CTCCAGCCAG AGGCTGACCT 4380  
GAGGGCACT GTCTCAGAT GACACACCC AGGAGCACCC TAGGTGAGGG GTGAGGGCCC 4440  
CCTTATGTGA ACCTCTTGCC TCTTCTTTTC TCCATCAGA GTGGTTGGAT GGAGCCATTG 4500  
GCCTCTTTT CTTCAGCGGG CCGTTCAACC TCTCTGACC ATGTTGTCTG GCTGAGGAGC 4560  
TACTAGAAAA GCTGAGTGGA GTCTCTTTTC CAACAGGATG ATGCATTGTC TCAATTCTCA 4620  
GGGCTGGAAT GAGCGGCTG GTCCCCAGA AAGCTGGAGT GGGGTACAGA GTTCAGTTT 4680  
CCTCTCTGTT TACAGTCTCT TGACAGTCCC ACGCCCATCT GGAGTGGGAG CTGGGAGTTA 4740

GTGTTGGAGA AGAAACAACA AAAGCCAATT AGAACCCTA TTTTAAAAA GTGCTTACTG 4800  
 TGCACAGATA CTCTTCAAGC ACTGGACGTG GATTCTCTCT CTAGCCCTCA GCACCCCTGC 4860  
 GGTAGGAGTG CGCCCTTAC CCACITGTGA TGGGGTACAG AGGCACITGC TCTTCTGCAT 4920  
 GGTGTTCAAT AGCGTGGAG TTTTATTTAT CTCTTCAAACT TTTGTACAAG AGCTCATGGC 4980  
 TTGTCTTGGG CTTCGTGAT TAAACCAAG GAAATGGAAG CCATTCOCCT GTTGTCTCC 5040  
 TTAGTCTTGG TCATCAGAAC CTCACTGGT ACCATATAGA TCAAAAGCTT TGTAAACCACA 5100  
 TGAACCAATA AACTCTTCCA TCCCTTAAAG AATAGAATAG TTTGTCCCTC TCATGGGAAT 5160  
 TGGGCTGTAT GTATATTGTT CTTCCTCCTT AGAATTTAGA GATACAAGAG TTCTACTTAG 5220  
 AACTTTTCAT GGACACAATT TCCACAACCT TTCAGATGCT GATGTAGAGC TATTGGGAAA 5280  
 GAACTTCCAA ACTCAGGAAG TTTGCAGAGA GCAGACAGCT AGAGATAACT CGGGACCCAG 5340  
 AGTTGGTCCA CAGATGTTAG ATGTATCCTA GCTTTTAGCC ATAAACCACT CAAAGATTCA 5400  
 GCCCCAGAT CCCACAGTCA GAACTGAATC TGCCTTGTGT GGAAGCCAGC AGTGGCCTTG 5460  
 GGAAGGAAGC CATGGCTGTG GTTCAGAGAG GGTGGGCTGG CAAGCCACTT CCGGGGAAAA 5520  
 CTCCCTCCGC CCGAGGTTTC TTCTTCTCTT AAGGAGAGAT TGTCTCACC AACCCGCTGC 5580  
 CTTCATGCTG CCTTCAAGC TAGATCATGT TTGCTTGTCT TAGAGAATTA CTGCAAAATCA 5640  
 GCCCAGTGC TTGGCGATGC ATTTACAGAT TTCTAGGCC TCAGGGTTT GTAGAGTGTG 5700  
 AGCCCTGGTG GGCAGGTTG GGGGCTCTGT CTCTGTCTG ATGCTGCTTG TAATCCATTT 5760  
 GGTGTACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 286 Protein sequence:  
 Protein Accession #: NP\_570843.1

1 11 21 31 41 51  
 MPLKHYLLLL VGCQAWAGL AYHGCPSECT CSRASQVECT GARIVAVPTP LPWNAMSLQI 60  
 LNTHTLWNE SPFLNISALI ALRIEKNELS RITPGAPRNL GSLRYLSLAN NKLQVLP IGL 120  
 FQGLDLSLESL LLSNQQLQI QPAHFSQCSN LKELQLHGNH LEYIPDGAFD HLVGLTKLNL 180  
 GRNSLTHISP RVFQHLGNLQ VLRLYENRLT DIPMGTFDGL VNLQELALQQ NQIGLLSPGL 240  
 FNNHNLRL YLSNNHISQL PPSIFMQLPQ LNRLTLFGNS LKELSLGIFG FMPNRLRELWL 300  
 YDNHISLSPD NVFSNLRQLQ VLILSRNQIS FISPGAPNGL TELRELSLHT NALQDLGDNV 360  
 FRLANLQNI SLQNNLRQLQ PGNI FANVNG LMAIQLQNNQ LENLPLGIFD HLGKLCBLRL 420  
 YDNPRCDS ILPLRNLL NQPRLGTDIV PVCFSANVR GQSLIIINVN VAVPSVHVPE 480  
 VPSYPETPMY PDTPSYPDIT SVSSTELTS FVEDYDILT IQVTDDRVSVM GMTQAQSGLA 540  
 IAAIVIGIVA LACSLAACVG CCCCKRSQA VLMQMKAPNE C

Seq ID NO: 287 DNA sequence  
 Nucleic Acid Accession #: NM\_002362  
 Coding sequence: 1..954

1 11 21 31 41 51  
 ATGCTCTCTG AGCAGAAGAG TCAGCACTGC AAGCCTGAGG AAGGCGTGA GGCCTCAAGAA 60  
 GAGGCCCTGG GCCTGGTGGG TGCAAGGCT CCTACTACTG AGGAGCAGGA GGCTGCTGTG 120  
 TCCTCTCTCT CTCTCTGTT CCCTGGCACC CTGGAGGAAG TGCTGTCTGC TGAATCAGCA 180  
 GGTCTCTCCC AGAGTCTCTA GGGAGCCTCT GCCTTACCCA CTACCATCAG CTTCCTTCTG 240  
 TGGAGGCAAC CCAATGAGGG TTCCAGCAGC CAAGAAGAGG AGGGCCCAAG CACCTCGCCT 300  
 GACGCAAGT CCTTGTCCG AGAAGCACTC AGTAACAAGG TGGATGAGTT GGCTCATTTT 360  
 TGCTCCGCA AGTATCGAGC CAAGGAGCTG GTCACAAGG CAGAAATGCT GGAGAGAGTC 420  
 ATCAAAATT ACAAGCCTG CTTCCTGTG ATCTTCGCA AAGCCTCCGA GTCCCTGAAG 480  
 ATGATCTTTG GCATTGAGCT GAAGGAAGTG GACCCCGCCA GCAACACCTA CACCTTGTG 540  
 ACCTGCTGG GCCTTCTCTA TGATGGCTG CTGGGTAATA ATCAGATCTT TCCCAAGACA 600  
 GGCCTTCTGA TAATCGTCTT GGGCACAATT GCAATGGAGG GCGACAGCGC CTCTGAGGAG 660  
 GAAATCTGG AGGAGCTGGG TGTGATGGGG GTGATGATG GGAGGGAGCA CACTGTCTAT 720  
 GGGGAGCCCA GGAATCTCT CACCCAGAT TGGGTGCAAG AAACTACCT GGAGTACCGG 780  
 CAGGTACCCG CGAGTAATCC TGCGGCTAT GAGTTCCTGT GGGTCCCAAG GGCTCTGGCT 840  
 GAAACCGACT ATGTGAAGT CCTGGAGCAT GTGGTCAGGG TCAATGCAAG AGTTCGCAAT 900  
 GCCTACCCAT CCTCGCTGA AGCAGCTTTG TTAGAGGAGG AAGAGGGAGT CTGA

Seq ID NO: 288 Protein sequence:  
 Protein Accession #: NP\_002353.1

1 11 21 31 41 51  
 MSSEQKSQHC KPEEGVEAQE EALGLVGAQA PTTEBQEAHV SSSSPLVPGT LEEVPAESA 60  
 GPPQSPQAS ALPTTISFTC WRQPNBGSST QREEGPSTSP DAESLFREAL SNKVDELHP 120  
 LLRYRAKEL VTKAEMLEVR IKNYKRCFPV IPGKASESLK MIPGIDVKEV DPASNTYTLV 180  
 TGLGLSYDGL LGNNQIFPKT GLLIIVLGTI AMEQDSASEE EIWEELGVMG VYDGREHTVY 240  
 GEPRKLLTQD WVENYLEYR QVPGSNPARY EFLWGPRLA STSYKVLEH VVRVNARVRI 300  
 AYPRLREAL LEEEBGV

Seq ID NO: 289 DNA sequence  
 Nucleic Acid Accession #: NM\_002362  
 Coding sequence: 46..1344

1 11 21 31 41 51  
 CGGCGGCGCG GCCCTGGTGT GGTCCCCACT GCTCTGGGG GGGCCATGGA CGAGGCCGTG 60  
 GCGCACTGA AGCAGGCGCT TCCCTGTGTG GCCAGTGC GCACGGTCCA CGTGGAGGTG 120  
 CATCAGCGCG GCAGCAGCAC TGCAAGAAA GAAGACATAA ACCTGAGTGT TAGAAAGCTA 180  
 CTCACAGAC ATAATATTGT GTTTGGTGAT TACACATGGA CTGAGTTTGA TGAACCTTTT 240  
 TTGACCAAG ATTGCAGTC TGTGTCTATT ATTGACAGC AATTAAAGGT TAAAGACTCA 300  
 CAGCCCATCG ATTTGAGTGC ATGCACTGTT GCACCTCACA TTTTCCAGCT GAATGAAGAT 360  
 GCGCCAGCA GTGAAAATCT GGAGGAAGAG ACAGAAAACA TAATTGACAG AAATCACTGG 420  
 GTTCTACCTG CAGCTGAATT CCATGGGCTT TGGGACAGCT TGGTATACGA TGTGGAAGTC 480  
 AATCCCATC TCCTCGATTA TGTGATGACA ACTTTACTGT TTTGACAGCA GAACGTCAAC 540

AGCAACCTCA TCACCTGGAA CCGGGTGGTG CTGCTCCACG GTCTCTCTGG CACTGGAAAA 600  
 ACATCCCTGT GTAAAGCGTT AGCCAGAAA TTGACATTA GACTTTCAAG CAGGTACCGA 660  
 TATGCCCATT TAATTGAAAT AAACAGCCAC AGCCTCTTTT CTAAGTGGTT TTCGAAAAGT 720  
 GGCAGCTGGG TAACCAAGAT GTTTCAGAAG ATTGAGGATT TGATTGATGA TAAAGACGCC 780  
 CTGGTGTTCG TGCTGATTGA TGAGGTGGAG AGTCTCACAG CGCCCGGAAA TGCCTGCAGG 840  
 GCGGGCACCG AGCCATCAGA TGCCATCCGC GTGGTCAATG CTGTCTTGAC CCAAAATGAT 900  
 CAGATTAATA GGCATTCCAA TGTTGTGATT CTGACCACTT CTAACATCAC CGAGAAGATC 960  
 GACGTGGGCT TCSTGGACAG GGCTGACATC AAGCAGTACA TTGGGCCACC CTCTGCAGCA 1020  
 GGCATCTTCA AAATCTACCT CTCTGTGTTG GAAGAACTGA TGAAGTGTC GATCATATAC 1080  
 CCTCGCCAGC AGCTGCTGAC CCTCGAGAG CTAGAGATGA TTGGCTTCAT TGAACAACAC 1140  
 GTGTCAAAAT TGAGCCTTCT TTGAATGAC ATTTCAAGGA AGAGCGAGGG CCTCAGCGGC 1200  
 CGGCTCCTGA GAAAACTCCC CTCTCTGGCT CATGCGCTGT ATGCCAGGC CCCCAACGTC 1260  
 ACCATAGAGG GGTCTCTCCA GGCCCTGTCT CTGGCAGTGG ACAAGCAGTT TGAAGAGAGA 1320  
 AAGAAGCTTG CAGCTTACAT CTGATCCTGG GCTTCCCCAT CTGGTGCTTT TCCCATGGAG 1380  
 AACACACAA CAGTAAGTGA GGTGCCCCA CACAGCCGTC TCCAGGGGAA TCCCTTCTGC 1440  
 AAACCAACG TTACTTAGAC TGCAAGCTAG AAAGCCACCA AGGCCAGGCT TTGTTAAAG 1500  
 AAGTGATTTC TATTATATGT GTTTTAAAT GCATCTGAG AGACAAACAT CTGTCTATT 1560  
 TCACTGTTTG TAAAGATAA TTCAGATTGT TTGTCTCCTT GTGAAGAAC ATCGAAACCT 1620  
 GTTGTCTCCC AGCCCAACCC CAGTGGATGG GATGCATAAT GCCAGCAAGT TTTGTTTAA 1680  
 AGCAAAAAGG AAGATTAAT GCAGGTGTTA TAGAAGCCAG AAGAGAACT GTGTCAACCT 1740  
 AAAGAGAGAT ATAATCAGT CATTAAAAAT GCACACATTA CTCCAGGTGG AAGGTGGCAA 1800  
 TTGCTTCTG ATATCAGCTG GTTTGATTGA GTGCAAAAT GTTTTCAAGA CTATTTAATG 1860  
 GATGTAAAAA AGCCTATTTC TACATTATAC CAACTGAGAA AAAAATGGTC GGTAAAGTGT 1920  
 TCTTTCATAA TAAATAATCA AGACATGGTC CCATTGTCAG GAAAGTGCA GACTCTGAGT 1980  
 GTTCAGGAGA AACACATGCT GGACATCCCT TGTAAACCGG TATGGCGGCC CCTGCTATGC 2040  
 TGGGATGTTT CTGCCCCAGG TTTTGTGTTG GCAATAACCT TATCACATT CTAATGAGGA 2100  
 TTCACATTAA TATAATATAA AATAAATAGG TCAGTTACTG GTCTCTTCTT GCCGAATGTT 2160  
 ATGTTTTGCT TTTATCTCAC AGTAAATAA ATATAATTA AAA

Seq ID NO: 290 Protein sequence:  
 Protein Accession #: NP\_004228

1 11 21 31 41 51  
 MDEAVGLKQ ALPCVAESPT VHVEVHQRGS STAKKEDINL SVRKLNRHN IVFGDYTWTE 60  
 FDEPFLTRNV QSVSLIDTEL KVKDSQPIDL SACTVALHIP QLNEDGPSSE NLEETENII 120  
 AANHVLPAE EFHGLWDSLV YDVEVKSHLL DYVMITLLFS DKNVNSNLIIT WNRVVLHHP 180  
 PGTGTSLLCK ALAQKLITRL SSRYRYGQLI EINSLSLFSK WFSSEGLKVT KMFQKIQDLI 240  
 DDKDALVFLV IDEVESLTAA RNACRAGTEP SDAIRVNAV LTQIDQIKRH SNVVILTSN 300  
 ITEKIDVAFV DRADIKQYIG PPSAAAIPIKI YLSCLEELMK CQIYIPROQL LTLRELEMIG 360  
 FIENNVSRLS LLLNDISRKS EGLSGRVLRL LFLAHLALYV QAPTVTIEGF LQALSIAVDK 420  
 QPEERKKLAA YI

Seq ID NO: 291 DNA sequence  
 Nucleic Acid Accession #: NM\_002658.1  
 Coding sequence: 77-1372

1 11 21 31 41 51  
 GTCCCGCGAG CGCCGTCGCG CCTCTCTGCC GCAGGCCACC GAGGCCGCGG CGGTCTAGCG 60  
 CCCCAGACCTC GCCACCATGA GAGCCCTGCT GGCGCGCCTG CTCTCTGCGG TCCTGGTCTGT 120  
 GAGCGACTCC AAAGGCGAGCA ATGAACCTCA TCAAGTTCCA TCGAACTGTG ACTGCTCTAA 180  
 TGGAGGAACA TGTGTGTCCA ACAAGTACTT CTCCAACATT CACTGGTGCA ACTGCCCAAA 240  
 GAAATTCGGA GGGCAGCACT GTGAAATAGA TAAGTCAAAA ACCTGCTATG AGGGGAATGG 300  
 TCACTTTTCG CAGGAGAAAG CCAGCACTGA CACCATGGGC CGGCCCTGCC TGCCCTGGAA 360  
 CTCTGCCACT GTCTTTCAGC AAACGTACCA TGCCACAGA TCTGATGCTC TTCAGCTGGG 420  
 CCTGGGGAAA CATTAATTAT GCAGGAACCC AGACAAACCG AGGCGACCCCT GGTGCTATGT 480  
 GCAGTGGGCG CTAAGCGCGC TTGTCCAAGA GTGCATGGTG CATGACTGCG CAGATGGAAA 540  
 AAAGCCCTCC TCTCTCCAG AAGAATTAAA ATTTCACTGT GGCCAAAAGA CTCTGAGGCC 600  
 CCGCTTTAAG ATTATTGGGG GAGAATTAC CACCATCGAG AACCAAGCCCT GGTTCGCGGC 660  
 CATCTACAGG AGGCACCGGG GGGGCTCTGT CACCTACGTG TGTGAGGGA GCCTCATCAG 720  
 CCGCTGCTGG GTGATCAGCG CCACACACTG CTTCATTGAT TACCCAAAGA AGGAGGACTA 780  
 CATCGTCTAC CTGGGTGCGT CAAAGCTTAA CTCCAACAGC CAAGGGGAGA TGAAGTTTGA 840  
 GGTGGAAGAC CTCATCTTAC ACAAGGACTA CAGCGCTGAC ACGCTTGCTC ACCACAACGA 900  
 CATTGCTCTG CTGAAGATGC GTTCCAAGGA GGGCAGGTGT GCGCAGCCAT CCGGACTAT 960  
 ACAGACCATC TGCCCTGCCCT CGATGTATAA CGATCCCCAG TTGGGCACAA GCTGTGAGAT 1020  
 CACTGGCTTT GGAAAAGAGA ATTCTACCGA CTATCTCTAT CCGGAGCAGC TGAATAATGAC 1080  
 TGTGTGAAG CTGATTTCCC ACCGGGAGTG TCAGCAGCCC CACTACTACG GCTCTGAAGT 1140  
 CACCACCAA ATGCTATGTG CTGCTGACCC CCAATGGAAA ACAGATTCTT GCCAGGGAGA 1200  
 CTCAGGGGGA CCGCTCTGCT GTTCCCTCCA AGGCCGCTG ACTTTGACTG GAATTGTGAG 1260  
 CTGGGGCGGT GGATGTGCCC TGAAGGACAA GCCAGCGCTC TACACGAGAG TCTCACATT 1320  
 CTTACCTTGG ATCCGAGTG ACACCAAGGA AGAGAATGGC CTGGCCCTCT GAGGGTCCCC 1380  
 AGGGAGGAAA CGGGCACCAC CCGCTTCTT GCTGGTGTG ATTTTTCAG TAGAGTCTATC 1440  
 TCCATCAGCT GTAAGAAGAG ACTGGGAAGA TAGGCTCTGC ACAGATGGAT TTGCTGTGTG 1500  
 CACCACCAAG GTGACGACA ATAGCTTTAC CCTCACGAT AGGCCTGGGT GCTGGCTGCC 1560  
 CAGACCTCTT GGGCAGGATG GAGGGGTGGT CTGACTCAA CATGTTACTG ACCAGCAACT 1620  
 TGTCTTTTTC TGGCTGAAG CCTGCAGGAG TTAAGAAAGG CAGGCGATCT CCGTGTGATG 1680  
 GGCTCGAAGG GAGAGCCAGC TCCCCGACG GGTGGGCATT TGTGAGGCC ATGGTTGAGA 1740  
 AATGAATAAT TTCCAATAA GGAAGTGTAA GCAGCTGAGG TCTCTTGGAG GAGCTTAGCC 1800  
 AATGTGGGAG CAGCGTTTG GGGAGCAGAG ACACTAACGA CTTCAGGGCA GGGCTCTGAT 1860  
 ATTCCATGAA TGTATCAGGA AATATATATG TGTGTGATG TTTGCACACT TGTGTGTGG 1920  
 GCTGTGAGTG TAAGTGTGAG TAAGAGCTGG TGTCTGATTG TTAAGTCTAA ATATTCTCT 1980  
 AAACCTGTGT GACTGTGATG CCACACAGAG TGTCTTTCT GGAGAGGTTA TAGGTCACTC 2040  
 CTGGGGCCTC TTGGGTCCCC CACGTGACAG TGCCTGGGAA TGTACTTATT CTGCAGCATG 2100  
 ACCTGTGACC AGCACTGTCT CAGTTTCACT TTCACATAGA TGTCCCTTTC TTGGCCAGTT 2160  
 ATCCCTTCTT TTTAGCCTAG TTTATCCAAT CCTCAGTGGG TGGGGTGAGG ACCACTCTCT 2220  
 ACACGTGAATA TTTATTTTTC ACTATTTTTC TTATATTTT TGAATTTTTC AATAAAGTGT 2280

ATCAATAAAA TGTGATT TTT CTGA

Seq ID NO: 292 Protein sequence:

Protein Accession #: NP\_002649.1

5  
10  
15

1	11	21	31	41	51	
MRALLARLLL	CVLVVSDSKG	SNELHQVPSN	CDCLNGGTCV	SNKYFSNIHW	CNCPKPKGQQ	60
HCEIDKSKTC	YEGNGHPYRG	KASTDTMGRP	CLPWSATVL	QQTTHARRSD	ALQLGLGKHN	120
YCRNPDNRRR	PWCYVQVGLK	PLVQECMVHD	CADGKKPSSP	PEELKFKQCG	KTLLRPRPKII	180
GGFTTIEHQ	PWFALYRRH	RGSVTVVCG	GSLISPCWVI	SATKCFIDYP	KKEDYIVYLG	240
RSRLNSNTQG	EMKFEVENLI	LHKDYSADTL	AHHNDIALLK	IRSKBGRCAQ	PSRTIQTICL	300
PSMYNDPQFG	TSCEITGPGK	ENSTDYLYPE	QLKMTVVKLI	SHRECOQPHY	YGSEVTTKML	360
CAADPQWKTD	SCQDSSGGPL	VCSLQGRMTL	TGIVSWGRGC	ALKDKPGVYT	RVSHPLPWIR	420
SHTKENGLA	L					

Seq ID NO: 293 DNA sequence

Nucleic Acid Accession #: NM\_001498

Coding sequence: 93...2006

20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70

1	11	21	31	41	51	
GGCAGGAGGC	TGAGTGTCCG	TCTCGCGCCC	GGAAGCGGGC	GACCGCGGTC	AGCCCGGAGG	60
AGGAGGAGGA	GGAGGAGGAG	GAGGGGGGCG	CCATGGGGCT	GCTGTCCAG	GGCTGCGCGC	120
TGAGCTGGGA	GGAACCAAG	CGCCATGCGG	ACCAAGTCCG	GCGGCAOGGG	ATCCTCCAGT	180
TCCCTGCACAT	CTACCAAGCC	GTCAGGACC	GGCACAAGGA	CGTTCTCAAG	TGGGGCGATG	240
AGSTGGAATA	CATGTTGTGA	TCTTTTGATC	ATGAAAATAA	AAAAGTCCGG	TTGGTCCGTG	300
CTGGGGAGAA	AGTTCTTGAA	ACTCTGCAAG	AGAAGGGGGA	AAGGACAAAC	CCAAACCATC	360
CTACCCCTTG	GAGACGAGAG	TATGGGAGTT	ACATGATTGA	AGGGACACCA	GGACAGCCCT	420
ACGAGGAGAC	AATGTCGAG	TTCAATACAG	TTGAGGCCAA	CATGCGAAAA	CGCGGAGAGG	480
AGGCTACTTC	TATATTAGAA	GAAAATCAGG	CTCTTTGCAC	AATAACTTCA	TTTCCCGAGT	540
TAGGCTGTCC	TGGGTTTCCA	CTGCCGAGG	TCAAACCCAA	CCCAGTGGAA	GGAGGAGCCT	600
CCAAGTCCCT	CTTCTTTCCA	GATGAAGCAA	TAAACAAGCA	CCCTCGCTTC	AGTACCTTAA	660
CAAGAAATAT	CCGACATAGG	AGAGGAGAAA	AGSTTGTGAT	CAATGTACCA	ATATTTAAGG	720
ACAAGAAATC	ACCATCTCCA	TTTATAGAAA	CATTTACTGA	GGATGATGAA	GCTTCAAGGG	780
CTCTTAAGCC	GGATCATATT	TACATGGATG	CCATGGGATT	TGGAATGGGC	AATGTCTGTC	840
TCCAGGTGAC	ATTCCAAGAC	TGCAGTATAT	CTGAGGCCAG	ATACCTTTAT	GATCAGTTGG	900
CTACTATCTG	TCCAATTGTT	ATGGCTTTGA	GTGCTGCATC	TCCCTTTTAC	CGAGGCTATG	960
TGTCTAGCAT	TGATTGTGCG	TGGGGAGTGA	TTTCTGCATC	TGTAGATGAT	AGAACTCGGG	1020
AGGAGCGAGG	ACTGGAGCCA	TTGAAGAACA	ATAACTATAG	GATCAGTAAA	TCCCGATATG	1080
ACTCAATAGA	CAGCTATTTA	TCTAAGTGTG	GTGAGAAATA	TAATGACATC	GACTTGACGA	1140
TAGATAAAGA	GATCTACGAA	CAGCTGTTGC	AGGAAGGCAT	TGATCATCTC	CTGGCCGAGC	1200
ATGTTGCTCA	TCTCTTTATT	AGAGACCCAC	TGACACTGTT	TGAAGAGAAA	ATACACCTGG	1260
ATGATGCTAA	TGAGTCTGAC	CATTTTGAGA	ATATTCAGTC	CACAAATTGG	CAGACAATGA	1320
GATTTAAGCC	CCCTCCTCCA	AACCTCAGCA	TTGGATGGAG	AGTAGAATTT	CGACCCATGG	1380
AGGTGCAATT	AACAGACTTT	GAGAACTCTG	CCTATGTGGT	GTTTGTGGTA	CTGCTCACCA	1440
GAGTGATCCT	TCTCTACAAA	TTGGATTTC	TCATTCCACT	GTCAAAGGTT	GATGAGAAAC	1500
TGAAGTAGC	ACAGAAAGA	GATGCTGTCT	TGCAGGGAAT	GTTTTATTTT	AGGAAAGATA	1560
TTTGCAAGG	TGGCAATGCA	GTGGTGGATG	GTGTGGGCAA	GGCCGAGAAC	AGCACGGAGC	1620
TGCTGTCAGA	GGAGTACACC	CTCATGAGCA	TAGACACCAT	CATCAATGGG	AAGGAAGGTG	1680
TGTTTCTCTG	ACTGATCCCA	ATTTGAACT	CTTACCTTGA	AAACATGGAA	GTGGATGTGG	1740
ACACGAGATG	TAGTATTCTG	AACCTACCTA	AGCTAATTAA	GAAGAGAGCA	TCTGGAGAAC	1800
TAATGACAGT	TGCGAGATGG	ATGAGGGAGT	TTATCGCAAA	CCATCCTGAC	TACAAGCAAG	1860
ACAGTGTGAT	AACTGATGAA	ATGAATTATA	GCCTTATTTT	GAAGTGTAA	CAATTTGCAA	1920
ATGAATTATG	TGAATGCCCA	GAGTTACTTG	GATCAGCATT	TAGGAAAGTA	AAATATAGTG	1980
GAAGTAAAC	TGACTCATCC	AACCTAGCAT	TCTACAGAAA	GAAAATGCA	TTATTGACGA	2040
ACTGGCTACA	QVTFQACGIS	TCTCAGCCCG	TGTGTATAAT	ATGAAGACCA	AATGATAGAA	2100
CTGTACTGTT	TTCTGGGCCA	GTGAGCCAGA	AATTGATTAA	GGCTTTCTTT	GGTAGGTAAA	2160
TCTAGAGTTT	ATACAGTGTG	CATGTACATA	GTAAGTATT	TTTGATTAA	AATGTATTTT	2220
AATAACATAT	CTAAAGTCAT	CATGAACCTG	CTTGATACAT	TTTAAATCT	TACTCTGGAG	2280
CAACCTACTG	TCTAAGCAGT	TTGTAAATG	TACTGTGTAAT	TGTACAATAC	TTGCATTCCA	2340
GAGTTAAAT	GTTTACTGTA	AAATTTTGT	CTTTTAAAGA	CTACCTGGGA	CCTGATTAT	2400
TGAAATTTT	CTCTTAAAA	ACATTTCTC	TCGTTAATTT	TCCTTTGTCA	TTTCTTTTGT	2460
TGCTACATT	AAATCACTTG	AATCCATTGA	AAGTGCTTCA	AGGTAATCT	TGGGTTTCTA	2520
GCACCTTATC	TATGATGTTT	CTTTTGCAAT	TGGAATAATC	ACTTGGTCA	CTTGCCCCAA	2580
GCTTTCCCT	CTGAATAAAT	ACCCATTGAA	CTCTGAAAA	AAAAAATAA	AAAA	

Seq ID NO: 294 Protein sequence:

Protein Accession #: NP\_001489

75  
80  
85

1	11	21	31	41	51	
MGLLSQGSPL	SWEETKRHAD	HVRRHGILQF	LHIYHAVKDR	HKDVLKWGDE	VEYMLVSFDH	60
ENKVLRLVLS	GEKVLETQKE	KGERTNPNHP	TLWRPEYGSY	MIEGTPQGFY	GGTMSEPTNV	120
EANMRKRKE	ATSIIEENQA	LCTITSFPRL	GCPGPTLPEV	KPNPVEGGAS	KSLFFPDEAI	180
NKHPRFSLT	RNRIRRRGEK	VVINVPFKD	KNTSPSPFET	PTEDDEASRA	SKPDHYMDA	240
MGFGMNCCL	QVTFQACGIS	EARYLYDQLA	TICPIVMALS	AASPPYRGYV	SDIDCRWGI	300
SASVDDRTRE	ERGLEPLKNN	NYRISKSRYD	SIDSYLSKCG	EKYNDIDLTI	DKEIYEQLLQ	360
EGIDHLLAQH	VAHLFIRDPL	TLPEEKIHL	DANESDHPE	IQSTNWQTM	FKPPFPNSDI	420
GWRVEFRPME	VQLTDFENSA	YVVFVLLTR	VILSYKLDPL	IPLSKVDENM	KVAQRDAVL	480
QGMFYFRDI	CKGNAVVDG	CGKAQNSTEL	AAEYTLMSI	DTIINGKEGV	FPGLIPIINS	540
YLENMEVDVD	TRCSILNLYK	LIKKRAGSEL	MTVARWMREF	IANHPDYKQD	SVITDEMNYN	600
LIILKNQIAN	ELCECEPELLG	SAFRKVKYSG	SKTSSN			

## WO 02/086443

Seq ID NO: 295 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 247-816

PCT/US02/12476

```
5 1 11 21 31 41 51
| | | | |
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCAGGTGC 120
CCTAGGGGGC ACATTTCCTCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGGGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGC CGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTGATATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
ATGACAGGAC ATGCTATTCC ACCAGGCCAA TTGATTCTC AGATTGATGA CTTCACTGGT 480
15 TTGAGCAAGG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
CAACGAGAAA TTAATGCTGA TATAAAACGT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
20 AAGCGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
AAGCACCTTA AGAAGAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
CACACCCCAA ATGCATAATC TCGTTAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCTGGCAT ATGTTACCGA ATCAAAATAGC 960
CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCATTTTAT 1020
25 TTGAAAGGAT AACTTGTGTT TTGTTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A
```

Seq ID NO: 296 Protein sequence:  
Protein Accession #: Eos sequence

```
30 1 11 21 31 41 51
| | | | |
MTDKTEKVAV DPETVFKRPR ECDSPSYQKR ORMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR ETASSPKSQR 120
35 EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIEKAA RCMRRDFVKH 180
LKKKLKRLMI
```

Seq ID NO: 297 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 247-815

```
40 1 11 21 31 41 51
| | | | |
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCAGGTGC 120
CCTAGGGGGC ACATTTCCTCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGGGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGC CGTCTCTCC TCCAGCAAGG 240
GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTGATATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
45 AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAAGCTT 420
ATGACAGGAC ATGCTATTCC ACCAGGCCAA TTGATTCTC AGATTGATGA CTTCACTGGT 480
TTGAGCAAGG ATAGGATGAT GCAGAAACCT GGTAGCAATG CACCTGTGGG AGGAAACGTT 540
ACCAGCAGTT TCTCTGGAGA TGACCTAGAA TGCAGAGAAA CAGCCTCCTC TCCCAAAGC 600
CAACAAGAAA TTAATGCTGA TATAAAACGT AAATTAGTGA AGGAACTCCG ATGCGTTGGA 660
50 CAAAAATATG AAAAAATCTT CGAAATGCTT GAAGGAGTGC AAGGACCTAC TGCAGTCAGG 720
AAACGATTTT TTGAATCCAT CATCAAGGAA GCAGCAAGAT GTATGAGACG AGACTTTGTT 780
AAGCACCTTA AGAAGAACT GAAACGTATG ATTTGAGAAT ACTTGTCCCT GGAGGATTAT 840
CACACCCCAA ATGCATAATC TCATTAAATGA TTGAGGAGAG AAAAGGATCA GATTGCTGTT 900
TTCTACAATG GAGCAGGATA TTGCTGAAGT CTCTGGCAT ATGTTACCGA ATCAACTGGC 960
60 CTTCCAGAGG CTAAGAAATT TCTGTTAGTA AAAGATGTTT TTTTCCCAA AGCGTTTAT 1020
TTGAAAGGAT AACTTGTGTT TTGTTTATTT TGTATTCCCA CCTGTGCTGG TAGATATTAT 1080
TAACCCATTA GGTAAATACT ATTACAGTCG TGGTTTCTGC A
```

Seq ID NO: 298 Protein sequence:  
Protein Accession #: Eos sequence

```
65 1 11 21 31 41 51
| | | | |
MTDKTEKVAV DPETVFKRPR ECDSPSYQKR ORMALLARKQ GAGDSLIAGS AMSKEKKLMT 60
70 GHAIPPSQLD SQIDDFTFPS KDRMMQKPGS NAPVGGNVTS SPSGDDLECR ETASSPKSQ 120
EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIEKAA RCMRRDFVKH 180
LKKKLKRLMI
```

Seq ID NO: 299 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 247-815

```
75 1 11 21 31 41 51
| | | | |
AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60
GGCCAAACGG GATCGGTGCT TCTGGTGAGA CGCTCCCCA TGCACATCAC TCCAGGTGC 120
CCTAGGGGGC ACATTTCCTCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180
GGGAGGGGCC ACAACTTCAC TGCCATTTTG TGAGGTGCGC CGTCTCTCC TCCAGCAAGG 240
80 GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300
CCCAGGGAAT GTGACAGTCC TTGATATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360
AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420
TGACAGGACA TGCTATTCCA CCGAGCCAAT TGGATTCTCA GATTGATGAC TTCCTGCTT 480
```



WO 02/086443

PCT/US02/12476

TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540  
 CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600  
 AACAGAAAT TAATGCTGAT ATAAAACGTA AATTAGTGAA GGAACCTCGA TGCCTGGAC 660  
 AAAAAATGA AAAAACTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720  
 AACGATTTT TGAATCCATC ATCAAGGAAG CAGCAGATG TATGAGACGA GACTTTGTTA 780  
 AGCACCTTAA GAAGAAACTG AAACGTATGA TTTGAGAATA CTGTCCCTG GAGGATTATC 840  
 ACACCCCAA TGCATAATCT CATTAAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900  
 TCTACAAATG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGAA TCAACTGGCC 960  
 TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAA GCGTTTATT 1020  
 TGAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080  
 AACCCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 300 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKARKLMT 60  
 GHAIPPSQLD SQIDDPGTG9 KDRMMQKPGS NAPVGGNVTS SPFGDDLECR ETASPPK9QQ 120  
 EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGTAVRKR FPESIIEKAA ROMRRDPVKH 180  
 LKKKLKRM

Seq ID NO: 301 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 247-812

1 11 21 31 41 51  
 AGTGTTCGGC TGGGGCAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTCCC ATCCCCCTTG 60  
 GGCCAAACAG GATCGGTGCT TCTGGTGAGA CGCCTCCCA TGCACATCAC TCCAGGTGC 120  
 CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180  
 GGGAGGGCGC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CGTCTCTCC TCCAGCAAGG 240  
 GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300  
 CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360  
 AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGA AAAGAGCTTA 420  
 TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCACTGGTT 480  
 TCAGCAAGA TGGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAATGTTA 540  
 CCAGCAATTT CTCTGGAGAT GACCTAGAAT GCAGAGGAAT AGCCTCCTCT CCCAAAGGCC 600  
 AACCAAGAAAT TAATGCTGAT ATAAATGTG AAGTAGTGAA GGAAATCCGA TGCCTTGGAC 660  
 AATATAGAAA AATCTTCGAA ATGCTTGAAG GAGTGCAAGG ACCTACTGCA GTCAGGAAAC 720  
 GATTTTTTGA ATCCATCATC AAGGAAGCAG CAAGATGTAT GAGACGAGAC TTTGTTAAGC 780  
 ACCTTAAGAA GAAACTGAAA CGTATGATTT GAGAATACCT GTCCCTGGAG GATTATCACA 840  
 CCCCAAATGC ATATCTCAT TAATGATTGA GGAGAGAAAA GGATCAGATT GCTGTTTTCT 900  
 ACAAATGGAG AGGATATTGC TGAAGTCTCC TGGCATATGT TACCGAATCA ACTGGCCTTC 960  
 CAGAGGCTAA GAAATTTCTG TTAGTAAAG ATGTTCTTTT TCCCAAAGCG TTTTATTGTA 1020  
 AAGGATAACT TGTGTTTTTG TTAATTTGTA TTCCACCTG TGCTGSTAGA TATTATTAAAC 1080  
 CQATTAGGTA AATACTATTA CAGTCGTGGT TTCTGCA

Seq ID NO: 302 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MTDKTEKVV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIIAGS AMSKEKKLMT 60  
 GHAIPPSQLD SQIDDPGTG9 KDRMMQKPGS NAPVGGNVTS NFSGDDLECR GIASSPPK9QQ 120  
 EINADIKQVY VKELRCQVY EKIFEMLEGV QGPTAVRKR FPESIIEAAR CMRRDPVKHL 180  
 KKKLKRMI

Seq ID NO: 303 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 247-815

1 11 21 31 41 51  
 AGTGTTCGGC TGGGACAGGC ACGCTGTGGC TGGCTACTTC CCTTCCTTCC ATCCCCCTTG 60  
 GGCCAAACAG GATCGGTGCT TCTGGTGAGA CGTCTCCCA TGCACATCAC TCCAGATGC 120  
 CCTAGGGGGC ACATTTCCCA CAACTCCAG AGGGCAGGTT TCTAGAAAGT GCCACCAAGT 180  
 GGGAGGGCGC ACAACTTCAC TGCCATTTTG TGAGGTGCCG CGTCTCTCC TCCAGCAAGG 240  
 GAAACAATGA CCGATAAAAC AGAGAAGGTG GCTGTAGATC CTGAAACTGT GTTTAAACGT 300  
 CCCAGGGAAT GTGACAGTCC TTCGTATCAG AAAAGGCAGA GGATGGCCCT GTTGGCAAGG 360  
 AAACAAGGAG CAGGAGACAG CCTTATTGCA GGCTCTGCCA TGTCCAAAGC AAAGAGCTTA 420  
 TGACAGGACA TGCTATTCCA CCCAGCCAAT TGGATTCTCA GATTGATGAC TTCACTGGTT 480  
 TCAGCAAGA TAGGATGATG CAGAAACCTG GTAGCAATGC ACCTGTGGGA GGAAACGTTA 540  
 CCAGCAGTTT CTCTGGAGAT GACCTAGAAT GCAGAGAAAC AGCCTCCTCT CCCAAAGGCC 600  
 AACCAAGAAAT TAATGCTGAT ATAAAACGTA AATTAGTGAA GGAACCTCGA TGCCTTGGAC 660  
 AAAAAATGA AAAAACTTC GAAATGCTTG AAGGAGTGCA AGGACCTACT GCAGTCAGGA 720  
 AACGATTTT TGAATCCATC ATCAAGGAAG CAGCAAGATG TATGAGACGA GACTTTGTTA 780  
 AGCACCTTAA GAAGAAACTG AAACGTATGA TTTGAGAATA CTGTCCCTG GAGGATTATC 840  
 ACACCCCAA TGCATAATCT CGTTAATGAT TGAGGAGAGA AAAGGATCAG ATTGCTGTTT 900  
 TCTACAAATG AGCAGGATAT TGCTGAAGTC TCCTGGCATA TGTACCGAA TCAACTGGCC 960  
 TTCCAGAGGC TAAGAAATTT CTGTTAGTAA AAGATGTTCT TTTTCCCAA GCGTTTATT 1020  
 TGAAGGATA ACTTGTGTTT TGGTTATTTT GTATTCCAC CTGTGCTGGT AGATATTATT 1080  
 AACCCATTAG GTAAATACTA TTACAGTCGT GGTTCCTGCA

Seq ID NO: 304 Protein sequence:  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MTDKTEKRVAV DPETVFKRPR ECDSPSYQKR QRMALLARKQ GAGDSLIAGS AMSKAKKLMT 60  
 5 GHAIPPSQOLD SQIDDFTGFS KDRMMQKPGS NAPVGGNVT SPSGDDLECR ETASSPKSQO 120  
 EINADIKRKL VKELRCVGQK YEKIFEMLEG VQGPTAVRKR PFESIIEKAA RCMRRDFVKH 180  
 LKKKLKRM

Seq ID NO: 305 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 87-689

1 11 21 31 41 51  
 15 CGTGGAGGCA GCTAGCGCGA GGCTGGGGAG CGCTGAGCCG CGCGTGTGTC CCTGCGCTGC 60  
 CCAGACTAGC GAACAATACA GTCCAGGATGG CTAAAGGTGA CCCCAGAGAA CCAAGGGGCA 120  
 AGATGTCCCG TTAATGCCCTTC TTTGTGCAGA CATGCRAGAGA AGAACATAAG AAGAAAAACC 180  
 CAGAGGTCCC TGTCAATTTT GCGGAATTTT CCAAGAAGTG CTCTGAGAGG TGAAGACGCA 240  
 TGTCGGGAA AGAGAAATCT AAATTTGATG AAATGGCAAA GGCAGATAAA GTGCGCTATG 300  
 20 ATCGGGAAT GAAGGATTAT GGACCAAGTA AGGAGAGCAA GAAGAAGAGG GATCCTAATG 360  
 CTCGCCAAAG GCCACCGTCT GGATCTCTCC TGTCTGTTC AGAATTCGCG CCCAAGATCA 420  
 AATCCACAAA CCGCGGCATC TCTATTGGAG ACGTGGCAA AAGCTGGGT GAGATGTGGA 480  
 ATAATTTAAA TGACAGTGAA AAGCAGCCTT ACATCACTAA GCGGCAAGG CTGAAGGAGA 540  
 AGTATGAGAA GGATGTTGCT GACTATAAGT CGAAAGGAAA GTTTGATGGT GCAAAGGGTC 600  
 25 CTGCTAAAGT TGCCCGGAAA AAGGTGGAAG AGGAAGATGA AGAAGAGGAG GAGGAAGAAG 660  
 AGGAGGAGGA GGAGGAGGAG GATGAATAAA GAAACTGTTC ATCTGTCTCC TTGTGAATAC 720  
 TTAGAGTAGG GCGGCGCCGT AATTGACACA TCTCTATTG GAGAAGTGTG TGTGCGCTC 780  
 ATTAGTTTAA ATTACAAAT TTGATCACGA TCATATTGTA GTCTCTCAA GTGCTCTAGA 840  
 AATTGTCACT GGTTTACATG AAGTGGCCAT GGGTGTCTGG AGCACCCTGA AACTGTATCA 900  
 30 AAGTTGACAA TATTTCACAA CATTTTAAA ATGAAAAGGC ACTCTGTGT TCTCTCACT 960  
 CTGTGCACTT TGCTGTTGGT GTGCAAGGC ATTTAAAGAT GTTCTGGCA TTTTCTTTT 1020  
 ATTGTAAAG TGGTGGTAAC TATGGTTATT GGCTAGAAAT CCTGAGTTT CAACTGTATA 1080  
 TATCTATAGT TTGTAATAAG AACAAAACAA CCGAGACAAA CCCTGTATGC TCCTGTCTCG 1140  
 GCGTTGAGGC TGTGGGGAAG ATGCGCTTTG GGAGAGGCTG TAGCTCAGGG CGTGCACTGT 1200  
 35 GAGGCTGGAC CTGTTGACTC TGCAGGGGGC ATCCATTAG CTTCAGGTTG TCTGTTTCT 1260  
 GTATATAGTG ACATAGCATT CTGCTGCCAT CTTAGCTGTG GACAAAGGGG GGTGAGCTGG 1320  
 CATGAGAAAT TTTTGTGTTT TAAGTGGGTG AGTTTAAA CTGTTGTGTT TTAACAACAA 1380  
 TATAGAACTC TTCATTGTCA GCAAGCAAAA GAGTCACTGC ATCAATGAAA GTTCAAGAAC 1440  
 CTCTGTACT TAACACAGAT TCGCAACGTT CTGTTATTTT TTTGTATGT TTAGAATGCT 1500  
 40 GAAATGTTTT TGAAGTTAAA TAAACAGTAT TACATTTTTA AAACCTCTCT CTATTATAAC 1560  
 AGTCAATTTT TGACTCACAG CAGTGAACAA ACCCCCACTC CATTTGATTT GGAGACTGGC 1620  
 CTCCTATAAA TGTGGGTAGC TTCTTTTATT ACTCAGTGGC CAGCTCACTT AGGGCTGAGA 1680  
 TGAAGGAGG GGCTACTTGA AGCTACTGTG TGATTTTGTG TGTGTCTGAG TGGCATTGAG 1740  
 ATGAAGTCTG GAGGAGTTAG GAGAACGACA TAGGCAAGGT TCAGCAGCCT TCCAAGTAT 1800  
 45 AGGAAGTGG GTGATTAGGA CTGAGGCTAT CTAGSTTTAA CTTTGTGCC ACCCTCACCC 1860  
 CCTATTTTGT GGGGCCAAT GCATTGCTAA ACAGCAATTT CAGAGTGAT GGTGTGTCAA 1920  
 AAATTAAGGC CTATTGTTT TTCTCTTTCA CCCCTACCCC CGGTGCTCCT GGCACATATC 1980  
 ACATTATTTG TGGTGCCCAA CATTTGGGGT CTGAGCCTG CTGCTGTCT CCTGATGCC 2040  
 AGTGAGGGTA TGTGGGATGG GGTGGTGGGG TAGGGGACGG TATCCTTTT TGTCTCTAC 2100  
 50 TTGGAACAC CAAACACCCC AAGGAAGATG ATAGGCTCCA TCTTGGGCA CCTGAGCTAT 2160  
 AGGCGAGGCT AATGGAATCA ACCATTCTG AGCACTAAAT GTATCATGAA AAGTTGAATG 2220  
 GCGTGTCTAT AAGTTTAGCT CATTCAGTGG AAATGTAGAT TGATGTTCAA TGTAAACTG 2280  
 GAAGGAGCTT GGTTTGTGTG TCAGTGGTTA TATTAGTGGG TAGTGTAACT TTTTATCCAG 2340  
 GTTGGGTGTA GGGGAGATGG CCACAGTAGC AAGTGGTGAC ACTAAATACC ATTTTGAAGG 2400  
 55 CTGATGTGTA TATACATCAT TACTGTCCGT AGCAATGAAG GATACAGTAC TGTGTTGTGG 2460  
 GTGAGTGTG TATTGCCCCA GCATTAAATAT TTGGGTGTGT ATGTTGTAGG CTATGAAACA 2520  
 CGCAGGAGTG TTTTGTGCT ATTAATTTTA AGAGAAAGCA GCTTTTCTT AAAATTCAC 2580  
 GTTGAGAAAC TTGATGTCT GGAGGCGGTG TCCTCTCCGC CTGTCTGGT CTGATGAG 2640  
 TACGAGTTAT GGTCAAGCTG ACAGCCTGAT CTCTTATGTT TCCATAGCCA TTGCTCTCC 2700  
 CATCAGAACT GTTTGTCTCT AATGTGTTCC TCTAGTTCTA GAAATGACC ACTAATTTAA 2760  
 60 AAAACTCGGT TGTGAGGTTT GCCCAGAGGC ACTGTGTTCA GAATTTCCCC TCCTGCTTCA 2820  
 GCCATGTCTT TGTCACTTGG CATTTCAAGC TAAAGCTTTA GCTTCCCAAT TCGTGTATGT 2880  
 CTAGGCCAAG ATTCGGGAGC TGTGCCAGC CTCGTCAAT ATGGAAGAGA AACCACTGC 2940  
 GGTCAAAAGG GAGTGATTG TTAAGTGGTG CGCTCTATC TCATAACTAG ATGTACCAAC 3000  
 CAGGGAAGGG CCAAGGATGG AAAGGGGTAA CTTTGTGTCT TCCAAAGTAG CTAAGCAGAA 3060  
 65 GTGGGGGAGC AGTTTAGCCA GATGATCTTT GATTAGGCAA ACATTGAGTT TTAAGAGGC 3120  
 TGTCAAGTTG AGGCCACTTG GTCCATTAGC TGGGGCAGCA AGATCACTAC TCAAGCTTT 3180  
 CACACTGTGG CAAGATTGCT CTCTAGTGG AATAATGCC TAGTTCTCT GAGATGATGT 3240  
 AAGTGGCATG ATGTTACCTA AGGCTTAGGC TTAGCTTGAT TTCTGGGCC ACTGTCTGTG 3300  
 70 TTCTTAAGAT GCCAACCTGT TGCTTTTTTT TTTTTTTTCC CCCATTTAAA AGGATAGTAC 3360  
 CTACTCCCTC TAACCACTC ACCCATTTCT TGAATGACAT TTTATCCTTC GGAAAGAAC 3420  
 AGGCTGTGAT GTAGTGACTA TTGTCTGTGT CTCCTGTGTG TGTCTGTCT TGTCACAAT 3480  
 GTATTTGGGG ACGTTGGATG CATTCAATTT CTGTAATAAA G

Seq ID NO: 306 Protein sequence:  
 Protein Accession #: NP\_005333.1

1 11 21 31 41 51  
 80 MAKGDPPKPK GKMSAYAFV QTCRESHKKK NPEVPVNPFE FSKKCSERWK TMSGKEKSKP 60  
 DEMAKADKVR YDREMKDYGP AKGKKKKDP NAPKRPSPGF FLPCSEFRPK IKSTNPISIS 120  
 GDVAKKLGEW WNNLNDSERQ FYITKAALKL EYKEDVDADV KSKGKFDGAK GPAKVARKKV 180  
 EEEEEEEEEE EEEEEEEEEE

Seq ID NO: 307 DNA sequence  
 Nucleic Acid Accession #: NM\_022342  
 Coding sequence: 1..2178

1 11 21 31 41 51  
5 ATGGGTACTA GGAAGAAAGT TCATGCATT GTCCGTGTCA AACCCACCGA TGACTTTGCT 60  
CATGAAATGA TCAGATACGG AGATGACAAA AGAAGCATTTG ATATTCACTT AAAAAAGAC 120  
ATTCGGAGAG GAGTTGTCAA TAAACAACAG ACAGACTGGT CGTTTAAGTT GGATGGAGTT 180  
TTCACGATG CCTCCAGGA CTTGGTTTAT GAGACAGTTG CAAAGGATGT GGTTCCTCAG 240  
CCCTCGATG GCTATAATGG CACCATCATG TGTATGGGG AGACGGGAGC TGGCAGAGCA 300  
10 ACACCATGA TGGGGGCAAC TGAGAATTAC AAGCACCGGG GGATCCTCCC TCGTGCCCTG 360  
AGCAGGTTT TTAGGATGAT CGAAGAACGC CCCACACATG CCATCACTGT GCGTGTTCCT 420  
ACTTGGAAA TCTATAATGA GAGCCTGTTT GATCTCCTGT CCATCTGCC CTATGTTGGA 480  
CCTCAGTCA CACCAATGAC CATCGTGGAA AACCTCAAG GAGTCTTCAT TAAGGGCTTG 540  
CAGTTCACC TCACAATGCA GGAGGAGGAT GCATTGAGCC TCCTTTTGA GGGTGAGACC 600  
15 ACAGGATTA TAGCTCCCA CACTATGAAC AAAAACTCTT CCAGATCACA CTGCATTTTC 660  
CCATCTACT TAGAGGCCCA TTCCCGGACC TTATCAGAGG AAAAGTACAT CACTTCCAAA 720  
TTAAGCTGG TGGATCTGGC AGGCTCAGAG AGGCTGGGA AGTCTGGTC TGAGGGCCAA 780  
TCCTGAAGG AAGCCACCTA CATCAACAAA TCGCTCTCAT TCCTGGAGCA GGCCATCATT 840  
CCCTTGGGG ACCAGAAGCG GGACCATATC CCCTTTCGGC AGTGCAAGCT CACCCAGCT 900  
20 TGAAGGACT CGTTAGGGGG AAACCTGCAAT ATGGTCTCG TGACAAACAT CTATGGAGAA 960  
CTGCCAGT TAGAAGAAAC GCTATCTTCA CTGAGATTG CCAGCAGGAT GAAGCTAGTC 1020  
CCAATGAGC CTGCCATCAA TGAAAAGTAT GATGCTGAGA GAATGGTCAA GAACCTGGAG 1080  
AGGAAGTAG CACTACTCAA GCAGGAGCTG GCTATCCATG ACAGCTGAC CAACCGCACC 1140  
TGTGACCT ATGACCCCAT GGATGAAATC CAGATTGCTG AGATCAACTC CCAGGTGGGG 1200  
25 GTTACCTCG AGGGGACACT GGACGAGATC GACATAATCA GCCTTAGACA GATCAAGGAG 1260  
TGTTCACCC AGTTCCGGGT GGTTCGAGC CACACAGAAC AGGAAGTGA GTCCACTTTG 1320  
GCAGGAAGT ACAGCTCAT TGACAGGAAT GACTTTGCG CCATTTCTGC TATCCAGAA 1380  
CGGGGCTTG TGGATGTTGA TGGCCACCTA GTGGGTGAGC CTGAAGGACA AAATTTTGA 1440  
TCGGAGTGG CCCCTTTCTC TACCAAACTT GGAAGAAAG CCAAGTCCAA GAAGACATTC 1500  
30 AAGAGCCAC TCAGGCCCGA CACCCACCC TCCAAACCG TGGCCTTTGA GGAGTTTAA 1560  
ATGAGCAAG GTAGTGAGAT CAACCGAAT TTCAAAGAAA ACAATCCAT CTTGAATGAA 1620  
GGAGGAAA GGGCCAGCGA GACCACACAG CACATCAATG CCATCAAGCG GGAGATTGAT 1680  
TGACCAAGG AGGCCCTGAA TTTCCAGAA TCACTACGGG AGAAGCAAG CAAGTACGAA 1740  
ACAAGGGG TGATGATCAT CGATGAGGAA GAATTCCTGC TGATCCTCAA GCTCAAAGAC 1800  
35 TCAAGAAGC AGTACCGCAG CGAGTACCAG GACCTGCGTG ACCTCAGGGC TGAGATCCAG 1860  
ATTGCGCAG ACCTAGTGA TCAGTGTGCG CACCGCTGCG TCATGGAATT TGACATCTGG 1920  
ACAATGAGT CCTTTGTCAT CCTGAGGAC ATGCAGATGG CACTGAAGCC AGGCGGCAGC 1980  
TCCGGCCAG GCATGTGCCC TGTGAACAGG ATTGTGCTCT TGGGAGAAGA TGACCAAGAC 2040  
AATTCAGCC AGCTGCAGCA GAGGTGCTT CCTGAGGCGC CTGATCCAT CTCCTTCTAC 2100  
40 ATGCCAAG TCAAGATAGA GCAGAAGCAT AATTACTTGA AAACCATGAT GGGCCTCCAG 2160  
AGGCACATA GAAAATAG

Seq ID NO: 308 Protein sequence:  
Protein Accession #: NP\_071737

45 1 11 21 31 41 51  
MGTRKKVHAF VRVKPTDDFA HEMIRYGGDK RSIDHLEKDD IRRGVVNQQ TDWSFKLDGV 60  
LHDASQDLVY ETVAQDVVSQ ALDGYNGTIM CYGQTGAGKT YTMGATENY KHRGILPRAL 120  
50 QQVFRMIEER PTHAITVRVS YLEIYNESLF DLLSTLPYVG PSVTMTIVE NPQGVFIKGL 180  
SVHLTSQRED AFSLLFEGET NRILASHTMN KNSSRSHCIF TIYLEAHSRT LSEKVIITSK 240  
INLVDLAGE SLGKSGSEGG VLKBYTYINK SLSPLEQAI ALGDQKRDHI PFRQCKLTHA 300  
LKDSLGGNEN MVLVTNIYGE AAQLEETLSS LRPAARMKLV TTEPAINEKY DAERMVNKLE 360  
KELALLKQEL AIHDLSTNRT FVTYDFMDEI QIAEINSQVR RYLEGTLDEI DIISLRQIKE 420  
55 VFNQFRVVL SQQEQVESTL RRYKTLIDRN DPAAISAIQK AGLVDVDGHL VGEPEGQNF 480  
LGVAPFSTKP GKAKSKKTP KEPLRPDTPP SKPVAPEEFK NEQGSSEINRI FKENSILNE 540  
RRKRASETTQ HNAIKREID VTKEALNFQK SLRERQKYE NKGLMIIDE EFLILILKLD 600  
LKQVRSYEQ DLRLDRAEIQ YCQLVVDQCR HRLMEFDIN YNESFVIED MQMALKPGGS 660  
IRPMVFPVNR IYSLGEDDQD KFSQLQQRVL PEGPDSISFY NAKVKIEQKH NYLKTMMGLQ 720  
QAHRK

Seq ID NO: 309 DNA sequence  
Nucleic Acid Accession #: CAT cluster

65 1 11 21 31 41 51  
TTTTTTTTTT TTTTTTTTAA TGCGTCTGT CATGCTCTGT CTACCAGGGT GAATTTCCAA 60  
AAATTTCTGC ATAGCAATTT TAGCCAAAAC TATATATGTT CTGGGGAGGA TAGGCATAGG 120  
CACATTGAAG ACCAAAGGAA AGAGTGAAGA AGTGTAGTTG GGTCTTGTG AATGGATGTT 180  
70 TAGATTGTCA AGAAAGTGG GCCAGAGGCC CCACCTCACA CTAGGACGGC AATGCTCTCT 240  
CATTAGTATC TCAGGCACCA TGGGTCTTAT TTGGTGTCTA AAGAAACACC CTCACAAAG 300  
TAATGAACCC TCAGCCTCCA GCTTCTCTTC TTCGGATTTC TTCTTAGGGC CTCCTTTTC 360  
CTTTTATGTT TCCAGTACCC TGAATTTCTT ATTCCCATCC CCATTAAAA TCTGCTTCAA 420  
AGAAAAACA AGAAGGACAC ATTCACITTA AGATCCAAAT GAATGATAAG AGCTTAAAC 480  
75 ATTATACTTA TCAGTATTAT TTGCATTTTT ATAGAAACCA AAACCATATT TCAACAAC

Seq ID NO: 310 DNA sequence  
Nucleic Acid Accession #: NM\_018622.2  
Coding sequence: 1-1140

80 1 11 21 31 41 51  
ATGGCGTGGC GAGGCTGGGC GCAGAGAGGC TGGGGCTGGG GCCAGGCGTG GGGTGGGTGG 60  
GTGGGGGGCC GCAGCTGCGA GGAGCTCACT GCGGTCTTAA CCGCGCCGCA GCTCCTCGGA 120  
85 CGCAGGTTTA ACTTCTTAT TCAACAAAAA TCGGGATTCA GAAAGCACC CAGGAAGGTT 180  
GAACCTCGAA GATCAGACCC AGGGCAAGT GGTGAAGCAT ACAAGAGAAG TGCTTTGATT 240  
CCTCCTGTGG AAGAAACAGT CTTTATCCT TCTCCTATC CTATAAGGAG TCTCATAAAA 300  
CCTTTATTTT TTACTGTGCG GTTACAGGC TGTGCATTG GATCAGCTGC TATTTGCCAA 360

TATGAATCAC TGAAATCCAG GGTCCAGAGT TATTTTGATG GTATAAAGC TGATTGGTTG 420  
 GATAGCATAA GACCACAAA AGAAGGAGAC TTCAGAAAGG AGATTAAACAA GTGGTGGAAAT 480  
 AACCTAAGTG ATGGCCAGCG GACTGTGACA GGTATTATAG CTGCAAAATGT CCTTGTATTG 540  
 TGTTTATGGA GAGTACCTTC TCTGCAGCGG ACAATGATCA GATATTTTAC ATCGAATCCA 600  
 GCCTCAAAGG TCCTTTGTTC TCCAATGTTG CTGTCAACAT TCAGTCACIT CTCCTTATTT 660  
 CACATGGCAG CAAATATGTA TGTTTTGTGG AGCTTCTCTT CCAGCATAGT GAACATTCTG 720  
 GGTCAAGAGC AGTTTCATGGC AGTGTACCTA TCTGCAGGTG TTATTTCCAA TTTTGTCACT 780  
 TACCTGGGTA AAGTTGCCAC AGGAAGATAT GGACCATCAC TTGGTGCACT TGGTGCCATC 840  
 ATGACAGTCC TCGCAGCTGT CTGCACTAAG ATCCCAGAAG GAGGCGTTCG CATTATTTTC 900  
 CTTCOGATGT TCAGTTTCAC AGCAGGGAAT GCCCTGAAAG CCATTATCGC CATGGATACA 960  
 GCAGGAATGA TCCTGGGATG GAAATTTTTT GATCATGCGG CACATCTTGG GGGAGCTCTT 1020  
 TTTGGAATAT GGTATGTTAC TTACGTCAT GAACGTGATT GGAAGAACAG GGAGCGGCTA 1080  
 GTGAAATCT GGCATGAAAT AAGGACTAAT GGGCCCAAAA AAGGAGGTGG CTCTAAGTAA

Seq ID NO: 311 Protein sequence:  
 Protein Accession #: NP\_061092.2

1 11 21 31 41 51  
 MAWRGWAQRG WSCQQAQWAS VGGRSCEELT AVLTPPOLLG RRFNFFIQK CGFRKAPRKV 60  
 EPRRSDPGTS GEAYKRSALI PFVEETVFYP SPYPIRELIK PLFPTVGFQT CAFGSAAIWQ 120  
 YESLKSQRVS YFDGIKADWL DSIRPQKEGD FRKEINKWVN NLSDDGQRTVT GIIAANVLVF 180  
 CLMRVPSLQR TMIRYFTSNP ASKVLCSPLM LSTFHSFSLF HMAANMYVLW SPSSSIIVNL 240  
 GGEQFMAYVL SAGVISNFVS YLGKVTAGRY GPSLGSAGAI MTVLAAVCTK IPEGRLAIFP 300  
 LMFPTFTAGN ALKAIIAMDT AGMILGWKFF DHAHLGGLAL FGIWYVTYGH ELIWKNNREPL 360  
 VKIWHIEIRTN GPKKGGGSK

Seq ID NO: 312 DNA sequence  
 Nucleic Acid Accession #: NM\_000625  
 Coding sequence: 195..3656

1 11 21 31 41 51  
 CTCTCGGCCA CCTTTGATGA GGGGACTGGG CAGTTCTAGA CAGTCCCGAA GTTCTCAAGG 60  
 CACAGGTCTC TTCTGTGTTT GACTGTCTCT ACCCGGGGGA GGCACTGCAG CCAGCTGCAG 120  
 GCCCCACAGT GAAGAACATC TGAGCTCAAA TCCAGATAAG TGACATAAGT GACCTGCTTT 180  
 GTAAAGCCAT AGAGATGGCC TGTCTTGGGA AATTTCTGTT CRAAGACCAA TTCCACCACT 240  
 ATGCAATGAA TGGGAAAAAA GGCACTCAACA ACAATGTGGA GAAAGCCCCC TGTGCCACCT 300  
 CAGTCCAGT GACACAGGAT GACCTTCAGT ATCACAACCT CAGCAAGCAG CAGAATGAGT 360  
 CCGCGCAGCC CCTCTGGGAS ACGGGAAAGA AGTCTCCAGA ATCTCTGGTC AAGCTGGATG 420  
 CAACCCCATT GTCTCTCCCA CGGCATGTGA GGATCAAAAA CTGGGGCAGC GGGATGACTT 480  
 TCCAAGACAC ACTTCACCAT AAGGCCAAAG GSATTTTAACT TTGCAGGTCC AAATCTTGCC 540  
 TGGGTCCAT TATGACTCCC AAAAGTTTGA CAGAGGAGCC CAGGGACAGC CCTACCCCTC 600  
 CAGATGAGCT TTACCTCAAA GCTATCGAAT TTGTCAACCA ATATTACGSC TCCCTCAAG 660  
 AGGCAAAAT AGAGGAACAT CTGCCCAGGG TGGAGCGGT AACAAAGGAG ATAGAAACAA 720  
 CAGTAACCTA CCAACTGACG GGAGATGAGC TCATCTTCGC CACCAAGCAG GCCTGGCGCA 780  
 ATGCCCCCAG CTGCATTGGG AGGATCCAGT GGTCCAACCT GCAGGTCTTC GATGCCCGCA 840  
 GCTGTTCCAC TGCCCGGGAA ATGTTTGAAC ACATCTGCAG ACACGTGCGT TACTCCACCA 900  
 ACAATGGCAA CATCAGGTG GGCATCACCG TGTTCGCCCA GCGGAGTGAT GGCAGACAG 960  
 ACTTCCGGGT GTGGAATGCT CAGCTCATCC GCTATGCTGG CTACCAAGAT CCAGATGGCA 1020  
 GCATCAGAGG AGTCCCGGAC AACGTGGAAT TCCTCAGCT GTGCTCGAG CTGGGCTGGA 1080  
 AGCCCAAGTA CGGCGCTTC GATGTGGTCC CCTGTGCTCT GCAGGCGAAT GGCCTGAGC 1140  
 CTGAGCTCTT CGAAATCCCA CTTGACCTTG TGCTTGAGGT GGCCTGGAA CATCCCAAT 1200  
 ACAGTGGTT TCGGGAACCT GAGCTAAAGT GGTACGCCCT GCCTGCAGTG GCCAACATGC 1260  
 TGCTTGAGGT GGGCGGCTTG GAGTTCACAG GTGTGCCCTT CAATGGCTGG TACATGGGCA 1320  
 CAGAGATCGG AGTCCCGGAC TTCTGTGATG TCCAGCGCTA CAACATCTCT GAGGAAGTGG 1380  
 GCAGGAGAA GGGCCTGGAA ACGCAACAAG TGGCTCTGCT CTGGAAGAAC CAGGCTGTCTG 1440  
 TTGAGATCAA CATTGCTGTG CTCCATAGTT TCCAGAAGCA GAATGTGACC ATCATGAGCC 1500  
 ACCACTCGGC TGCAGAACTC TTCTGAAGT ACATGCAGAA TGAATACCGG TCCCGTGGGG 1560  
 GCTGCCCGGC AGACTGGATT TGGCTGGTCC CTCCATGTC TGGGAGCATC ACCCCGCTGT 1620  
 TTCACCAGGA GATGCTGAAC TAAGTCTGT CCCCCTTCTA CTACTATCAG GTAGAGGCCT 1680  
 GGAACACCA TGCTGGCAG GACGAGAAGC GGAGACCCAA GAGAAGAGAG ATTCCATTGA 1740  
 AAGTCTTGGT CAAAGCTGTG CTCTTTGCCT GTATGCTGAT GCGCAAGACA ATGGCGTCCC 1800  
 GAGTCAGAGT CACCATCTCT TTTGCGACAG AGACAGGAAA ATCAGAGGCG CTGGCTGGG 1860  
 ACCTGGGGGC CTATTTCAGC TGTGCCCTCA ACCCAAGGT TGTCTGCATG GATAAGTACA 1920  
 GGCTGAGCTG CCTGGAGGAG GAACGGCTGC TGTGGTGGT GACCAGTACG TTTGGCAATG 1980  
 GAGACTGCCC TGGCAATGGA GAGAAACTGA AGAAATCGCT CTTCATGCTG AAAGAGCTCA 2040  
 ACACAAATTT CAGGTACGCT GTGTTTGGCC TGGGCTCCAG CATGTACCCCT CGGTTCTGCG 2100  
 CCTTGTCTCA TGACATTGAT CAGAAGCTGT CCGACCTGGG GGCTCTCAG CTCACCCCGA 2160  
 TGGGAGAAGG GGATGAGCTC AGTGGGCAAG AGGAGCGCTT CCGCAGCTGG GCGCTGCAAA 2220  
 CCTTCAAGGC AGCCTGTGAG ACCTTTGATG TCCGAGGCAA ACAGCACATT CAGATCCCA 2280  
 AGCTCTACAC CTCCATATGT ACCTGGGACC CGCACCATA CAGGCTCGTG CAGGACTCAC 2340  
 AGCCTTTGGA CTTCAGCAAA GCCCTCAGCA GCATGCATGC CAAGAACGTG TTCACCATGA 2400  
 GGCTCAAACT TCGGCAAAAT CTACAAAGTC CGACATCCAG CCGTGCCACC ATCCTGGTGG 2460  
 AACTCTCTCT GAGGATGGC CAAGGCTGA ACTACCTGCC GGGGGAGCAG CTGGGGTTT 2520  
 GCCCAGGCAA CCAGCCGCCC CTGCTCCAG GCATCTGGA GCGAGTGGTG GATGGCCCCA 2580  
 CACCCACCA GGCAGTGCCT CTGAGGCCCC TGGATGAGAG TGGCAGCTAC TGGGTCACTG 2640  
 ACAAGAGGCT GCCCCCTGTC TCACTCAGCC AGGCCCTCAC CTACTTCTG GACATACCA 2700  
 CACCCCAAC CCAGCTGTCT CTCCAAAGC TGGCCCAAGT GGCCACAGAA GAGCCTGAGA 2760  
 GACAGAGCT GAGAGCGCTG TGCCAGCCCT CAGAGTACAG CAAGTGGAAG TTCACCAACA 2820  
 GCCCCACATT CTGAGAGGTG CTAGAGGAGT TCCGCTCCT GCGGTGTCT GCTGGCTTCC 2880  
 TGCTTTCCCA GCTCCCATTT CTGAAGCCCA GGTCTACTIC CATCAGCTCC CCCCAGGATC 2940  
 ACACGCCAC GAGATCCAC CTGACTGTGG CCGTGGTCC CTACCAACCC CGAGATGGCC 3000  
 AGGGTCCCT GCACACGCG GTCTGCAGCA CATGGCTCAA CAGCTGAAG CCCCAGACC 3060  
 CAGTGCCTG CTTTGTGCGG AATGCCAGCG GCTTCCACCT CCCCAGGAT CCCCCTCATC 3120

CTTCGATCCT CATCGGGCTT GGCACAGGCA TCGCGCCCTT CGCAGTTTC TGGCAGCAAC 3180  
 GGCTCCATGA CTCCAGCAC AAGGAGTGC GGGGAGGCCG CATGACCTTG GTGTTTGGGT 3240  
 GCGCGCGCCC AGATGAGGAC CACATCTACC AGGAGGAGAT GCTGGAGATG GCCCAGAAGG 3300  
 GGGTGTGTGA TGGGTGTGAC ACAGCCTATT CCGCCTGCC TGGCAAGCCC AAGGTCTATG 3360  
 TTGAGGACAT CCGTGGGACG CAGCTGGCCA GCGAGGTGCT CCGTGTGCTC CACAAGGAGC 3420  
 CAGGCCACCT CTATGTTTGC GGGGATGTGC GCATGGCCCG GGACGTGGCC CACACCTGA 3480  
 AGCAGCTGGT GGTGCCCAAG CTGAAATTGA ATGAGGAGCA GGTGAGGAC TATTTCTTTC 3540  
 AGCTCAAGAG CCAAGAGGCG TATCAGAGG ATATCTTTGG TGCTGTATT CTCTACGAGG 3600  
 CGAAGAAGGA CAGGTGTGCG GTGAGCCCA GCAGCCTGGA GATGTCAGCG CTCTGAGGGC 3660  
 CTACAGGAGG GGTAAAGCT GCGGACAG AACTTAAGGA TGGAGCCAGC TCTGCATTAT 3720  
 CTGAGGTAC AGGGCTGGG GAGATGAGG AAGTGATAT CCCCAGCCT CAAGTCTTAT 3780  
 TTCTCAACG TTGCTCCCA TCAAGCCCTT TACTTGACCT CTAACAAGT AGCACCTGG 3840  
 ATTGATCGGA GCCTC

Seq ID NO: 313 Protein sequence:  
 Protein Accession #: NP\_000616

1 11 21 31 41 51  
 MACPWKFLFK TKFHOYAMNG EKGINNVEK APCATSSPVT QDDLQYHNLS KQNESQPPL 60  
 VETGKKSPEP LVKLDPATPLS SPREVRIKNW GSGMTFQDIL HHKAKGILTC RSKSCLG SIM 120  
 TPKSLTRGPR DKPTPPDELL PQAIEFVNQY YGSLKEAKIE EHLARVEAVT KEIETTVTYQ 180  
 LTGDELIPAT KQAWRNAPRC IGRIQWSNLQ VFDARSCSTA REMFEHICRH VRYSTNNGNI 240  
 RSAITVFPQR SDGKHDFRVW NAQLIRYAGY QMPDGSIRGD PANVEFTQLC IDLGWPKPYG 300  
 RFDVVLVLQ ANGRDPELFE IPPDLVLEVA MEHPKYEWFR ELELKWYALP AVANMLLEV 360  
 GLEFPFCFPN GWYMGTEIGV RDPDQVORYN ILEEVGRMRG LETHKLASLN KDQAVVEINI 420  
 AVLESFQKQN VTIMDHSAE ESFMYMQNE YRSRGCCPAD WIVLVPPMSG SITPVPHQEM 480  
 LNYVLSPPFY YQVEANKTHV WQDEKRRPKR REIPLKVLVK AVLFACMLMR KTMASRVVVT 540  
 ILPATETGKS EALAWDLGAL FSCAPNPXVY CMDKYRLSCL EERLLLVVT STFGNGDCPG 600  
 NGELKLLSLP MLKELNNKFR YAVFGLGSSM YPRFCAPAH IDQKLSHLGA SOLTPMGEGD 660  
 ELGGQEDAFR SNAVQTFKAA CSTFDVVRGKQ HIQIPKLYTS NVTWDPHHYR LVQDSQPLDL 720  
 SKALSSMEAK NVFTMRLLSR QNLQSPSSR ATILVELSCE DGQGLNLYPG EHLGVCPGNO 780  
 PALVQGILER VVDGPTFHQA VRLEALDESG SYWVSDKRLP PCSLSQALTY FLDITTPPTQ 840  
 LLLQKLQVVA TESEPRORLE ALQCPSEYSK WKPTNSPTFL EVLEFPFSLR VSAGFLLSQL 900  
 PILKPRFVSI SSPRHTPTE IHLTVAVVTY HTRDGQGPLH HGVCSTWLSN LKPDQPVPCF 960  
 VRNAGFHLPL EDPSEPCILLI GPQGTIAPFR SFWQRLHDS QHKGVRGGRM TLVFGCRRPD 1020  
 EDHYQEMEL EMAQGVVLAH VHTAYSRLPG KPVVYVDIL RQQLASEVLR VLHEKPGHLY 1080  
 VCGDVRMARD VAHTLKQLVA AKLKLNEBQV EDYFPQLKSQ KRYHEDIFGA VFPYEAKRDR 1140  
 VAVQPSSELM SAL

Seq ID NO: 314 DNA sequence  
 Nucleic Acid Accession #: XM\_087254  
 Coding sequence: 47..2332

1 11 21 31 41 51  
 AGAGTACGTG TTTACAGATA AAACCTGGTAC ACTGACAGAA AATGAGATGC AGTTTCGGGA 60  
 ATGTTCAATT AATGGCATGA AATACCAAGA AATTAATGGT AGACTTGATC CCGAAGGACC 120  
 AACACAGACG TCTTCAGAGG GAAACTTATC TTATCTTAGT AGTTTATCCC ATCTTAACAA 180  
 CTATATCCAT CTTCACACCA GTTCTCTTTT CAGAACCACT CCGTAAAATG AAACCTGAAT 240  
 AATTAAGAGA CATGATCTCT TCTTTAAAGC AGTCAGTCTC TGTCACACTG TACAGATTAG 300  
 CAATGTTCAA ACTGACTGCA CTGGTGTATG TCCCTGGCAA TCCAACTGG CACCATCGCA 360  
 GTTGGAGTAC TATGCATCTT CACCAGATGA AAAGGCTCTA GTAGAAGCTG CTGCAAGGAT 420  
 TGGTATTGTG TTTATTGGCA ATTCTGAAGA AACTATGGAG GTTAAAACTC TTGGAAAACT 480  
 GGAACGGTAC AAACCTGCTC ATATCTTGGA ATTTGATTCA GATCGTAGGA GAATGAGTGT 540  
 AATTGTTTCA GACCTCTCAG GTGAGAAGTT ATTATTGTCT AAAGGAGCTG AGTCATCAAT 600  
 TCTCCCTAAA TGTATAGGTG GAGAAATAGA AAAAACCAGA ATTCATGTAG ATGAATTGTC 660  
 TTTGAAAGGG CTAAGAACCT TGTGTATAGC ATATAGAAAA TTTACATCAA AAGAGTATGA 720  
 GGAATATAGT AAACGCATAT TTGAAGCCAG GACTGCCTTG CAGCAGCGGG AAGAGAAATT 780  
 GGCAGCTGTT TTCCAGTTCA TAGAGAAAGA CCTGATATTA CTGGAGGCA CAGCAGTAGA 840  
 AGACAGACTA CAAGATAAAG TTCGAGAAAC TATTGAAGCA TTGAGAATGG CTGGTATCAA 900  
 AGTATGGGTA CTACTGGGGG ATAAACATGA AACAGCTGTT AGTGTGAGTT TATCATGTGG 960  
 CCATTTTCAT AGAACCATGA ACATCCTTGA ACTTATAAAC CAGAAATCAG ACAGCGAGTG 1020  
 TGCTGAACAA TTGAGGCAGC TTGCCAGAGG AATTACAGAG GATCATGTGA TTCAGCATGG 1080  
 GCTGTAGTGT GATGGGACCA GCCTATCTCT TGCACTCAGG GAGCATGAAA AACTATTTAT 1140  
 GGAAGTTTGC AGAATTTGTT CAGCTGTATT ATGCTGTGTT ATGGCTCCAC TGCAGAAAGC 1200  
 AAAAGTAATA AGACTAATAA AAATATCACC TGAGAAACCT ATAACATTGG CTGTGGGTGA 1260  
 TGGTGTCTAT GACGTAAGCA TGATACAAGA AGCCCATOTT GGCAATAGGA TCATGGGTAA 1320  
 AGAAGGAAGA CAGGCTGCAA GAAACAGTGA CTATGCAATA GCCAGATTGA AGTTCTCTCT 1380  
 CAATATGCTT TTTGTTTCAT GTCATTTTGA TTATATTAGA ATAGCTACCC TTGTACAGTA 1440  
 TTTTTTTCAT AAGAATGTGT GCCTTATCAC ACCCCAGTTT TTATATCAGT TCTACTGTTT 1500  
 GTTTTCTCAG CAACATTGTT ATGACAGCGT GTACCTGACT TTATACAATA TTTGTTTAC 1560  
 TTCCCTACCT ATTCTGATAT ATAGTCTTTT GGAACAGCAT GTAGACCCCTC ATGTGTTTAC 1620  
 AAATAAGCCC ACCCTTTATC GAGACATTAG TAAAAACCGC CTCTTAAGTA TTAACAATT 1680  
 TCTTTATTGG ACCATCTTGG GCTTCAGTCA TGCTTTTATT TTCTTTTTC GATCCTATT 1740  
 ACTAATAGGG AAGATATACAT CTCTGCTTGG AAATGGCCAG ATGTTTGGAA ACTGGACATT 1800  
 TGGCATTGTT GTCTTCACAG TCATGGTTAT TACAGTCACA GTAAAGATGG CTCTGGAAAC 1860  
 TCATTTTGGG ACTTGGATCA ACCATCTOGT TACCTGGGGA TCTATTATAT TTTATTTTGT 1920  
 ATTTTCCCTG TTTTATGGAG GGATTTCTCT GCCATTTTGG GGCTCCCGA ATATGATTTT 1980  
 TGTGTTTATT CAGCTCTCTG CAAGTGGTTC TGCTTGGTTT GCCATAATCC TCATGGTTGT 2040  
 TACATGTCTA TTTCTTGATA TCATAAGAA GGTCTTTGAC CGACACCTCC ACCCTACAAG 2100  
 TACTGAAAG GCACAGCTTA CTGAAACAAA TGCAGGTATC AAGTGCTTGG ACTCCATGTG 2160  
 CTGTTTCCCG CAAGGAGAGG CAGCGTGTGC ATCTGTGGGA AGAATGCTGG AACGAGTTAT 2220  
 AGGAAGATGT AGTCCAACCC ACATCAGCAG ATCATGGAGT GCATCGGATC CTTTCTATAC 2280  
 CAACGACAGG AGCATCTTGA CTCTCTCCAC AATGGACTCA TCTACTTGTT AAAGGGGCG 2340

WO 02/086443

PCT/US02/12476

5

10

15

20

25

30

35

40

45

50

55

60

TAGTACTTTG TGGGAGCCAG TTCACCTCCT TTCCTAAAAT TCAGTGTGAT CACCCCTGTTA 2400  
 ATGGCCACAC TAGCTCTGAA ATTAATTTCCT AAAATCTTTG TAGTAGTTCA TACCCACTCA 2450  
 GAGTTATAAT GGCACACAAA CAGAAAGCAT TAGTACAAGC CCCTCCCAAC ACCCTTAATT 2520  
 TGAATCTGAA CATGTTAAAA TTGAGAAATA AAGAGACATT TTTCATCTCT TTGCTCTGGT 2580  
 TGTCCCTTGT GCTTATGGGA CTCTAAATGG CATTTCACTC TGTGTCTGAG GCCATTATAT 2640  
 TTTAATATAA ATGTAGAAAA AAGAGAGAAA TCTTAGTAAA GAGTATTTTT TAGTATTAGC 2700  
 TTGATTATTG ACTCTTCTAT TTAATCTCCT TTCTGTAAAT TATGCTGAAA GTTTGCTTTG 2760  
 AGAACTCTAT TTTTATTATA GAGTTATATT TAAAGCTTTT CATGGGAAAA GTTAATGTGA 2820  
 ATACTGAGGA ATTTTGTGCT CTCAGTGACC TGTGTTGTTA ATTCATTAAAT GCATTCTGAG 2880  
 TTCACAGAGC AAATTAGGAG AATCATTTCC AACCATTAAT TACTGCAGTA TGGGGAGTAA 2940  
 ATTTATACCA ATTCCTCTAA CTGTACTGTA ACACAGCCTG TAAAGTTAGC CATATAAATG 3000  
 CAAGGGTATA TCATATATAC AAATCAGGAA TCAGGTCCGT TCACCGAACT TCAAAATTGAT 3060  
 GTTTACTAAT ATTTTGTGTA CAGAGTATAA AGACCCATA GTGGGTAAAT TAGATACTAT 3120  
 TAGCATATTA TTAATTTAAT GTCTTTATCA TTGGATCTTT TGCGATGCTT AATCTGGTTA 3180  
 ACATATTATA ATTTGCTTTT TTCTCTTTA CCTGAAGGCT CTGTGTATAG TATTTTCATGA 3240  
 CATGTTGTTA CAGTTTAACT ATATCAATAA AAAGTTTGGG CAGTATTAA ATATTGCAAA 3300  
 TATGTTTAAAT TATACAAATC AGAATAGTAT GGGTAATTA ATGAATACAA AAAGAGAGC 3360  
 CTCCTTCTCG AGCCGACTTA GACATGCTCT TCCCTTTCTA TAAGCTAGAT TTTAGATAAA 3420  
 AGGGTTTCAG TTAATAATCT TATTTTCAGG TTATGTCATC TAACTATAG CAAACTACCA 3480  
 CAATACAGTG AGTTCTGCCA GTGTCCCACT ACAAGGCATA TTTCAAGGTG GGTGTGGGAA 3540  
 TGTAAAAAATG CTCAACTTGT ATCAGGTAAT GTTAGCAATA AATTAAATGC TAAGAAATGAT 3600  
 TAATCGGGTA CATGTTACTG TAAATTAECT ATTGCACCTC AAAACCTAAC TTCCATCTCTG 3660  
 AATTATACAA GTAGTTCACT ATTGTCAITT GTTTTGTGTT TATTGAAAAG TAATGTTGTC 3720  
 TTAAGATTTA GAAGTGATTA TTAGCTTGAG AACTATTACC CAGCTCTAAG CAAATAATGA 3780  
 TTGTATACAT ATTAAGATAA TGGTTAAATG CGGTTTACC AAGTTTCCC TTGAAAATGT 3840  
 AATTCTCTTA TGGAGATTTA TTGTGACGCC CTAAAGCTCC TCCCAATTTC ATGAATATAA 3900  
 GGCTTCTAGA ATTTGACTCG CAGGGGAAAG AATGCTAGAG ACAGAAATTA AGACTTTATC 3960  
 CTGTTTGTCT TGTAACTAT TATTTTCTTG CTAATGTAACT ATTGTGCTG TCCAGTGATG 4020  
 TAAGGATATT AAGTTATTTA GCTAAATATT AATTTTCAAA AATAGTCCCT CTTTAACTTA 4080  
 GATATTTCAAT AGCTGGATT TGGAAAGATCT GTTATCTGAG AAGTACTAAA AAGAATAATA 4140  
 CAACGTACAA TGCTCTGCAAT CACTAATTCA TGTTCAGAAA GAGGAAATAA TGAAGATATA 4200  
 CTCAGTAGAG TACTAGTGGG GAGGATATGG AAATTTGCTC ATAAATCTC TTATAAAACG 4260  
 TGCAATAAAC AAAATGACAC CCAGTAGGCC TGCAATTACAT TTACATGACC GTGTTTATTT 4320  
 GCCATCAAT AAACCTAGTA CTGACACCGA ACAAGACTC CAAAGTCATA AATAGCCTA 4380  
 TGACCAACTG CAGCAAGACA GGAGGTCAGC TCCCTTATAA TGGTGCTTAA AGTGTGATTG 4440  
 ATGTAATTTT CTGTACTCAC CATTGAAGT TAGTTAAGGA GAACCTTATT TTTTAAAAAA 4500  
 AAGTAAATGG CAACCACTAG TGTGCTCATC CTGAACCTGT ACTCCAAATC CACTCCGTTT 4560  
 TTAAGCAAAA ATTATCTTGT GATTTTAAGA AAAGAGTTT CTATTTTATT AAGAAAGTAA 4620  
 CAATGCACTG TCGAAGCTTT CAGTAGTTT CTAGTGCTAT ATTCACTCTG TAAACTCTT 4680  
 ACTACGTAAC CAGTAATCAC AAGGAAAGTG TCCCTTTTGC ATATTCTTT AAAATCTTT 4740  
 CTTTGGAAAG TATGATGTTG ATAATTAAC TACCCTTATC TGCCAAAACC AGAGCAAAAT 4800  
 GCTAAATACG TTATTTCTAA TCAGTGGTCT CAAATCGATT TGCTCCCTT TGCTCTGCTT 4860  
 GAGGCGTGTA AGCCTGAAGA TAGTGGCAAG CACCAAGTCA GTTTCAAAAT TTGCCCCCTCA 4920  
 GCTGCTTAA GTGACTCAGC ACCCTGCTCT AGCTTCAGCA GGCGTAGGCT CACCCCTGGGC 4980  
 GGAGCAAGT ATGGGCCAGG GAGAACTACA GCTACGAAGA CCTGCTGTG AGTTGAGAAA 5040  
 AGGGGAGAA TTAGGTCTG AATTTTCTAA CTGTCTCTT TCTTGGTCT AAAGCTCATA 5100  
 ATACACAAGG GCTTCAGAC CTGAGCCACA CCCAGGCCCT ATCTGAACA GGAGACTAAA 5160  
 CAGAGGCAAA TCAACCTTAG GAAATACCTG CATTCTGCCC TACGGTTAGT ACCAGGACTG 5220  
 AGGTCAATTC TACTGGAAA GATTTGTAGA TTGAACCTAT CTGATCGCTT GAGACTCTTA 5280  
 ATAGGCAGGA GTCAAGGCCA CTAGAAAAT GACAGTTAAG AGCCAAAAGT TTTTAAAAA 5340  
 TGCTACTCTG AAAAATCTCG TGAAGGCTGT AGGAAAAGGG AGAATCTTCC ATGTTGGTGT 5400  
 TTTTCTGTA AAGATCAGT TGGGGTATGA TATAAGCAGG TATTATTA AATAACACAC 5460  
 CAAAGAGTTA CGTAAACAT GTTTTATTA TTTTGGTCCC CAGTACAGA CATTTTATT 5520  
 CTATTTTGA ATGAGTTATC TATTTTATA AAAGTAAAC ACTATTAAAG TGCTGTTTAA 5580  
 TGTAAATAA CTGAATGT GTTCTATAA AAAATAGATC ATAACTCATG ATATGTTTGT 5640  
 AATCATGTTA TTTAGATT TTAGAGGAA TGAGTACTCTG GAAATATTGT AGCAATACTT 5700  
 GGTTTAAAT TTTGGACCTG AGACACTGTG GCTGTCTAAT GTAATCTTT AAAAATCTCT 5760  
 TGCAATGTCA GTAAATGTAG TATATTATTG TACAGCTACT CATAATTTTT TAAAGTTTAT 5820  
 GAAGTTATAT TTATCAAAATA AAAACTTTCC TATAT

Seq ID NO: 315 Protein sequence:  
 Protein Accession #: XP\_087254

65

70

75

80

1 11 21 31 41 51  
 MQPRECSING MKYQEIINGRL VPEGPTPDSS EGNLSYLSSL SHLNNLSHLT TSSSFRTSPE 60  
 NETELIKEED LFFKAVSLCH TVQISNVQTD CTGDGPWQSN LAPSQLEYA SSPDEKALVE 120  
 AAARIGIVFI GNSBETMEVK TLGLKERYKL LHILFSDSR RRMSVIVQAP SGEKILFAK 180  
 AESSILPKCI GGEIEKTRIH VDEFALKGLR TLCIAYRKFT SKEYEIDKR IFEARTALQQ 240  
 REEKLAAPVQ FIEKDLILLG ATAVEDRLQD KVRETTIEALR MAGIKVWVLT GDKHETAVSV 300  
 SLSCGFHRT MNILELINQK SDSECAQLR QLARRITEDH VIQHGLVVDG TSLSLALREH 360  
 EKLFMEVCRN CSAVLCCRMA PLQKAKVIRL IKISPEKPIT LAVGDGANDV SMIOEAVGI 420  
 GIMGKEGRQA ARNSDYAIAR PKFLSKLLFV HGFFYYIRIA TLVQVFFYKN VCFITPQRLY 480  
 QFYCLFSQQT LYDSVYLTLV NICFTSLPIL IYSLLEQHVD PHVLQNKPTL YRDISKNRLL 540  
 SIKTFPLYNTI LGFSHAFIFP FGSYLLIGKD TSLGNGQMF GNWTFGLTVF TVMVIITVVK 600  
 MALETHFWTV INHLFWGSI IFYFVPSLFY GGLWPFPLGS QNMVFPVPLQ LSSGSAAWFAI 660  
 ILMVVTCLFL DIKIKVFDRL LHPTSTTEKAQ LTETNAGIKC LDSMCCFPPEG EAACASVGRM 720  
 LERVIGRCSP THISRWSAS DPFTTNDRSI LTLSTNDSST C

Seq ID NO: 316 DNA sequence  
 Nucleic Acid Accession #: NM\_004473  
 Coding sequence: 661..1791

85

1 11 21 31 41 51  
 CTCGCCACGC GTCCGCGGGG CTGAGACCC ACGCCGTGGA GAGGACCAGC CTCAGGTGCG 60

CCGCGCTGGG CCGCGGCCCC GACCTCGCTG CCCCGGCTC GCTCTCTGCG CGTGGGCGCT 120  
TACCGCCACC TTGGCTCGG GGGCAGGGCA TGGGGGGGCC CGCCAGATC GCCCAGCGCC 180  
AGTACTAACT GCGCTCGCTC TGGCTTCTGA GCCCGAAGCC TCTCTGCGCG GCACAACTTA 240  
GGCAGTATC CTAACTAGC GGGCACCACA GACCACTGCG ACCCAACCCA ACCCAGGAT 300  
CACTTCCGGA CCGCTCGACC GCGCGGCACC AGCGCGCAAG GGAACCTTCA GCCCGAGACC 360  
AGAGTCCAGT CCGCGTCCGC AGGCCACCGC CGTGCCTGCG CTCGAGAAGC ACAACGCGGG 420  
CTGAGCGGTC GGCTAGCGGG TCACTCCCGA GCCTCTGTCT GCACCGCGCC AGCCCCAGAC 480  
CAGGACGCT GAGCCTCCAG CGCGCGCCAG CCTGGGCGCG TGGGCTCTCC GGGCCAGGCC 540  
GGCAGATCC CTTGAGCTCT CCGCAGAAGG GCGAGCGCTC CGTTCGGGG AGCCAGGCC 600  
CGCCCCGCC CCGGACAGC CGCGGGGATC CAGAGCCCGG GGGTGGCGGA CGCCCGCGCC 660  
ATGACTGCCG AGAGCGGGCC GCGCGCGCG CAGCGGAGG TGCTGGCTAC CGTGAAGGAA 720  
GAGCGCGCGG GCGGCAAGCG CCGCTGCGAG CGCGGGAAGC CGCCCTACAG CTACATCGCG 780  
GGCGGCGCGG CCGGCAAGCG CCGCTGCGAG CGCGGGAAGC CGCCCTACAG CTACATCGCG 840  
CTCATCGCCA TGGCGCATCG GCACGCGGCC GAGCGCGCG TCACTGCTGG CGGCATCTAC 900  
AAGTTTATCA CCGAGCGCTT CCGCTTCTAC CGCAGCAACC CCAAAAAGTG GCAGAACAGC 960  
ATCCGCCACA ACCTCACACT CAACGACTGC TTCTCAAGA TCCGCGCGGA GCGCGCGCG 1020  
CCGGTAAGG GCACTACTG GCGCTTCCAG CCAACGCGG AGGACATGTT CGAGAGCGGC 1080  
AGCTTCTGCG CGCGCGCGCA GCGCTTCAAG CGCTCGGACC TCTCCACCTA CCGGCTTAC 1140  
ATGACGAGCG CGCGCGCTGC CGCAGCGGCC GCTGCGCGAG CGCGCGCGCG CGCGCGCGCC 1200  
CGCGCATCT TCCAGGCGCG GGTGCGCGCC GCGCGCGGCC CTACCGCGGG CGCGCTCTAT 1260  
GCAGGCTACG CGCGCGCTGC GCTGCGCGCG CGCGCTCAG TCTACTACC CGCGCGCTCG 1320  
CCCGGCCCTT CCGCGCTCTT CGGCTTGGTT CTTGAGCGCG CGCTCAGGCC AGAGCTGGGG 1380  
CCCGCACGCT CGCGCGCGCG CGGCTTCTGC GCTTCTGGCT CGCGCGCGCG CCGCGCTACC 1440  
ACCACGCGCT ACCAGCGCGC AGGCTGACCC GGGCGCGGCC CGGCCAACCC CTCTGCTTAT 1500  
CGCGCTGCTT ACAGCGCGCC GAGCGCGCG TACCGCGAG GCGCGCGCG TGGGATCTTT 1560  
GCGCTGCTG GCGCGCTGCG GGGACCGGCT TCGCGCGCG CGGGCGCGAG CAGTGGCGCG 1620  
GTGAGACCA CGTGGAGCTT CTACCGCGCG ACCTGCGCG GCGAGTTCGG AGCGCTGGGA 1680  
GCCTGTACA ACCCTGGCGG GCACTCGGA GGGGCGAGT CAGGCGCTTA CCATGCTGCG 1740  
CATGCTGCG CTATCTCGGG TGGATAGAT CGGTTCGTGT CCGCATGTG AGCCAGCGTA 1800  
GGGAGAAAA CTATAGACA CATCGGCTGT TCACAGCTTC CCGCAACCT GAGAACGAAC 1860  
AGGAATGAG AGAGGACTCA ACTGGGACCC AGTGGAAAA GACCGAGCAG GCCACAGAGG 1920  
CTCGTCTCC CCGGACAGC CGTAGGACC CTGTGTACTC TGTAAAGGG AGGAGGTGGG 1980  
CGGAGGACG CAGAGCCCTT GCACTGGCAC AGGACCCCTC GATGGAGCGA AGCCCTCAAA 2040  
CGGATGCTT TCTGCACTT TATCGGGAG GGTCTTGGC GGTAACAGA GGGCAGCGTA 2100  
GTGTCAACAC CAGAGACAGG GATCCAAATT GTGGGGAATC AGTTTCAGCC TTCCATGTGC 2160  
TGCGGAACT CGGCGCTTCT TACGCGGTTT GTCTCTAGT GCCTTTAACT GCGTTACTAC 2220  
AATAAAGGC TGCGGACGCG CTTTCTTCT TAAAGTGGG AGGACAAATT TGCAAAAGAA 2280  
ATAGGCTTTT CTTCTTTT AAATGGAGA AATCTCTGCT CTGGTTGACC TGGGCTGGTT 2340  
TTCCCTGCT CTGAGAACTT GAGACCTAGC TCGAGTGA ACTGTGCTC AGCACTCCAG 2400  
TCCATCAC CAGAACTTCA GTCTCCCCA TCTGTTCAC TAGAGGGCTG CAGGACTCTA 2460  
TCCACGCGCC CCGGCTTATC ATTCAGGGCC CCATCATCTT GGATGCTGCC CTGCGTATT 2520  
GGCAGCAATG GTGGCCACC CAGGCGCTCT GAGTAGCCAC CCAAGGCTTA GCCGCTGTT 2580  
TAGGGAACG AAAAGAGTTC ATGCGCAAGC GTCTAACCTA AAGTCCAGG ATTGGCTCCA 2640  
GGCAGCAATT ATATCATAAC TTATTGAAT TTTAGCAGG ACCTGCTGCT AATTTCATGG 2700  
CTGTACTG CAGTCAATA ATCTGCTTTT CCATTATAAG GCAGAGAGAA GTACATTCGT 2760  
TCATTGATG AGAGATTCCA AGGAGCCCGT TGGTACGCC TTTAGTAGC TGGGAGGGG 2820  
TCTTCCATCC CAGCACCCCT CTGTACACC TCAGCAGCCT CCGCATGCA AAAAGGAAAG 2880  
AGAAAAATTA AGTTAGGCGA GTCAGTAAAG TGAGCTTTAG AAAGAACTG GAATTTTAA 2940  
TTCATTTTGT ATCTTGCTTA AGTAGCAGGC TCACTAAAA TAGAGAAAGT CCAATAACTC 3000  
TCCCGCTTTC CTTTGAGAAA TCTTTAAGTT TCGATTCTGG AGCAAAACTT TTCAGCTTA 3060  
AATATTTCAG AGGCTCCATT CACAGCTTTC AGATAAACTG GAGTGTTCAG ATGGAGTGT 3120  
TTAATAAAAA TCTTTGAGCA AGTGAGTTAT GGCAAGAGAA ACTCAGGCTC TTTCTGTATA 3180  
AACTTAACAG GGAAGGCGTG GGTGTGAAA AAGAAGATTG TATGAAAACC ATTGGTAATT 3240  
TTTATTTTTT ATTTTGGGA CTGCACTATC CTGTTACAGA AGACATGTGA ACTTGGTTCA 3300  
GTCCAAATGG GAATTGTAT AAACCAAGTC TCTCATTAG AATATGTTG CAAGCCACAT 3360  
ATGTAATTTT AATATTCTA GTAGCCACAT TAATAAGTN AAAAGAAACA AAAAAAATA 3420  
AA 3480

Seq ID NO: 317 Protein sequence:  
Protein Accession #: NP\_004464

1 11 21 31 41 51  
| | | | |  
65 FKHLTHYROI DTRANSCRIP TIONFACTOR TTFMTAESGP PPPQPEVLAT VKEERGETAA 60  
GAGVPGEATG RGAGGRRRKR FLQRGKPPYS YIALIAMAIA HAPERRLTIG GIYKPIITERF 120  
PFYRDNPKKW QNSIRHNLTL NDCFLKI PRE AGRPGKGNYN ALDPNAEDMF ESGSFLRRRK 180  
RFKRSDLSTY PAYMHDAAAA AAAAAAAAAA AAAAAIFPGA VPAARPPYPG AVYAGYAPPS 240  
70 LAAPPVYVYP AASPGRVRF GLVPERPLSP ELGPAPSGPG GSCAFASAGA PATTTGYQPA 300  
GCTGARPNP SAYAAAYAGP DGAYPQAGS AIFAAAGRLA GPASPPAGGS SGGVETTVD 360  
YGRTPGQFG ALGACYNPGG QLGGASAGAY HARHAAAYPG GIDRPFVSAM

Seq ID NO: 318 DNA sequence  
Nucleic Acid Accession #: NM\_005688  
Coding sequence: 126..4439

1 11 21 31 41 51  
| | | | |  
80 CCGGCGAGGT GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCCTGGAGC 60  
AGGGGCGCAG GAATTCGTAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCGCTCAG 120  
AGAGATGAA GGATATCGAC ATAGGAAAAG AGTATATCAT CCCCAGTCCT GGGTATAGAA 180  
GTGTAGCGGA GAGAACCAGC ACTTCTGGGA CGCAGAGAGA CGTGGAAGAT TCCAAGTTCA 240  
85 GGAGAACTCG ACCGTTGGAA TGCCAAGATG CCTTGGAAAC AGCAGCCCGA GCCGAGGGCC 300  
TCTCTCTTGA TGCTCTCAGC CATTCTCAGC TCAGAACTCT GGATGAGGAG CATCCCAAGG 360  
GAAAGTACCA TCATGGCTTG AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACCCAGC 420  
ACCCAGTGGG CAATGCTGGG CTTTTTCTCT GTATGACTTT TTCGTGGCTT TCTTCTCTGG 480



	CCCGTGTGGC	CCACAAGAAG	GGGGAGCTCT	CAATGGAAGA	CGTGTGGTCT	CTGTCCAAGC	540
	ACGAGTCTTC	TGACGTGAAC	TGCAGAGAC	TAGAGAGACT	GTGGCAAGAA	GAGCTGAATG	600
	AAGTTGGGCC	AGACGCTGCT	TCCCTGGGAA	GGGTGTGTG	GATCTTCTGC	CGCACCAGGC	660
5	TCATCCTGTC	CATCGTGTGC	CTGATGATCA	CGCAGCTGGC	TGGCTTCAGT	GGACCAGCCT	720
	TCATGGTGAA	ACACCTCTTG	GAGTATACCC	AGGCAACAGA	GTCTAACCTG	CAGTACAGCT	780
	TGTTGTAGT	GCTGGGCCTC	CTCCTGACGG	AAATCGTGCG	GTCTTGGTCG	CTTGCACTGA	840
	CTTGGGCATT	GAATTACCA	ACCGGTGTCC	GCTTGGGGGG	GGCCATCCTA	ACCATGGCAT	900
	TTAAGAAGAT	CCTTAAGTTA	AAGAACATTA	AAGAGAAATC	CCTGGGTGAG	CTCATCAACA	960
10	TTTGTCCAA	CGATGGGCAG	AGAATGTTTG	AGGCAGCAGC	CGTGGCAGC	CTGCTGGCTG	1020
	GAGGACCCGT	TGTTGCCATC	TTAGGCATGA	TTTATAATGT	AATTATTCTG	GGACCAACAG	1080
	GCTTCTCTGG	ATCAGCTGTT	TTTATCCTCT	TTTACCCAGC	AATGATGTTT	GCATCACGGC	1140
	TCACAGCATA	TTTCAGGAGA	AAATGCGTGG	CCGCCACGGA	TGAACGTGTC	CAGAAGATGA	1200
	ATGAAGTTCT	TACTTACATT	AAATTTATCA	AAATGTATGC	CTGGGTCAAA	GCATTTCTCT	1260
15	AGAGTGTCTA	AAAAATCCGC	GAGGAGGAGC	GTGATATT	GGAAAAAGCC	GGGTACTTCC	1320
	AGGGATATCA	TGTGGGTGTG	GCTCCCATG	TGGTGGTGAT	TGCCAGCGTG	GTGACCTTCT	1380
	CTGTTTATAT	GACCTTGGGC	TTGATCTGA	CAGCAGCACA	GGCTTTCACA	GTGGTGACAG	1440
	TCCTCAATTC	CATGACTTTT	GCTTTGAAAG	TAACACCGTT	TTCACTAAGG	TCCCTCTCAG	1500
	AAGCCTCAGT	GGCTGTTGAC	AGATTTAAGA	GTTTGTCTCT	AATGGAAGAG	GTTCACATGA	1560
20	TAAAGAACAA	ACCAGCCAGT	CCTCACATCA	AGATAGAGAT	GAATAATGCC	ACCTTGGCAT	1620
	GGGACTCCTC	CCACTCCAGT	ATCCAGAACT	CGCCCAAGCT	GACCCCAAAA	ATGAAAAAAG	1680
	ACAGAGGGGC	TTCCAGGGGC	AAGAAGAGA	AGGTGAGGCA	CTGCAGCGC	ACTGAGCATC	1740
	AGGCGGTGCT	GGCAGAGCAG	AAAGGCCACC	TCCTCCTGGA	CAGTGACGAG	CGGCCACGTC	1800
	CCGAAGAGGA	AGAAGGCAAG	CACATCCACC	TGGGCCACCT	GGCTTACAG	AGGACACTGC	1860
25	ACAGCATCGA	TCCTGAGATC	CAAGAGGGTA	AATCTGTTGG	AATCTGCGGC	AGTGTGGGAA	1920
	GTGGAAAAAC	CTCTCTCATT	TCAGCCATT	TAGGCCAGAT	GACGCTTCTA	GAGGGCAGCA	1980
	TTGCAATCAG	TGGAACCTTC	GCTTATGTGG	CCGACGAGGC	CTGATCTCTC	AATGCTACTC	2040
	TGAGAGACAA	CATCTGTTT	GGGAAGGAAT	ATGATGAAGA	AAGATACAAC	TCTGTGCTGA	2100
	ACAGCTGCTG	CTGAGGCTCT	GACCTGGCCA	TTCTTCCAG	CAGCGACCTG	ACGGAGATTG	2160
30	GAGAGCGAGG	AGCCAACTG	AGCGGTGGGC	AGCGCCAGAG	GATCAGCCTT	GCCCGGCTCT	2220
	TGTATAGTGA	CAGGAGCATC	TACATCCTGG	ACGACCCCTC	CAGTGCCTTA	GATGCCCATG	2280
	TGGGCAACCA	CATCTTCAAT	AGTGCTATCC	GGAAACATCT	CAAGTCCAAG	ACAGTTCTGT	2340
	TTGTATACCA	CCAGTTACAG	TACCTGGTTG	ACTGTGATGA	AGTGATCTTC	ATGAAGAGAG	2400
	GCTGTATTAC	GGAAAGAGGC	ACCCATGAGG	AACTGATGAA	TTTAAATGGT	GACTATGCTA	2460
35	CCATTTTAA	TAACCTGTG	CTGGAGAGAG	CACCGCCAGT	TGAGATCAAT	TCAAAAAAGG	2520
	AAACCACTGG	TTACAGAGAG	AGGTACAAAG	ACAGGGTCC	TAAACAGGA	TCAGTAAAGA	2580
	AGGAAAAAGC	AGTAAAGCCA	GAGGAAGGGC	AGCTTGTGCA	GCTGGAAGAG	AAAGGGCAGG	2640
	GTTCAGTGCC	CTGGTCAGTA	TATGTTGTCT	ACATCCAGGC	TGCTGGGGGC	CCCTTGGCAT	2700
	TCTCGGTTAT	TATGCGCTT	TTTATGCTGA	ATGTAGGAG	CACCGCCTTC	AGCACCTGGT	2760
40	GGTTGAGTGA	TGAGATCAAG	CAAGGAAGCG	GGAAACCCAC	TGTGACTCGA	GGGAACGAGA	2820
	CCTCGGTGAG	TGACAGCATG	AAGGACATC	CTCATATGCA	GTACTATGCC	AGCATCTACG	2880
	CCCTCTCCAT	GGCAGTCATG	CTGATCCTGA	AAGCCATTGC	AGGAGTTGTC	TTTGTCAAGG	2940
	GCACGCTGCG	AGCTTCTCTC	CGGCTGCATG	ACGAGCTTTC	CCGAAGGATC	CTTCGAAGCC	3000
	CTATGAAGTT	TTTTGACACG	ACCCCAACAG	GGAGGATTC	CAACAGGTTT	TCCAAAGACA	3060
45	TGGATGAAGT	TGACGTGCGG	CTGCCCTTCC	AGGCCGAGAT	GTTCATCCAG	AACGTTATCC	3120
	TGGTGTCTCT	CTGTGTGGGA	ATGATCGCAG	GAGTCTTCCC	GTGGTTCCTT	GTGGCAGTGG	3180
	GGCCCTCTGT	CATCCTCTTT	TCAGTCTCTG	ACATTGTCTC	CAGGGTCTCG	ATTCCGGGAGC	3240
	TGAAGCGTCT	GGACAATATC	ACGCACTCAC	CTTCTCTCTC	CCACATCAAG	TCCAGCATAC	3300
	AGGGCTCTGC	CACCATCCAC	GGCTACAATA	AAGGGCAGGA	GTTCCTGAC	AGATACCCAG	3360
50	AGCTGCTGGA	TGACAAACAA	GCTCTCTTTT	TTTTGTTTAC	GTGTGCGATG	CGGTGGCTGG	3420
	CTGTGCGGCT	GGACCTCATC	AGCATCGCCC	TCATCACCAC	CACGGGGCTG	ATGATGGTTC	3480
	TTATGCAACG	GCAGATTCCT	CCAGCCTATG	CGGGTCTCGC	CATCTCTTAT	GCTGTCCAGT	3540
	TAAAGGGGCT	GTTCAGTTT	ACGGTCAGAC	TGGCATCTGA	GACAGAAGCT	CGATTCACTC	3600
	CGGTGAGAGG	GATCAATCAG	TACATTAAAG	CTCTGCTCTT	GGAGGACACT	CCGAGAATTA	3660
55	AGAACAGAGC	TCCCTCCCTC	GACTGGCCCC	AGGAGGGAGA	GGTGACCTTT	GAGAACGCAG	3720
	AGATGAGGTA	CCGAGAAAC	CTCCCTCTTG	TCCTAAAGAA	AGTATCCTTC	ACGATCAAC	3780
	CTAAAGAGAA	GATTGGCATT	GTGGGGCGGA	CAGGATCAGG	GAAGTCTCTG	CTGGGGATGG	3840
	CCCTCTTCOG	CTGGTGGAG	TTAICTGGAG	GCTGCATCAA	GATTGATGGA	GTGAGAATCA	3900
	GTGATATTGG	CTTGGCGGAC	CTCGAAGCA	AACCTCTCTAT	CATTCTCTCA	GAGCCGGTGC	3960
60	GTTCAGTGG	CATGTCAGAG	TCAAATTTGG	ACCCCTTCAA	CCAGTACACT	GAGACCCAGA	4020
	TTTGGGATGC	CCTGGAGAGG	ACACACATGA	AAGAATGTAT	TGCTCAGCTA	CACTCTGAAC	4080
	TGAATCTGA	AGTGATGGAG	AATGGGGATA	ACTTCTCAGT	GGGGGAACCG	CAGCTCTTGT	4140
	GCATAGCTAG	AGCCCTGCTC	CGCCACTGTA	AGATTCTGAT	TTTAGATGAA	GCCACAGCTG	4200
	CCATGAGCAC	AGAGACAGAC	TTATTGATT	AAGAGACCAT	CCGAGAAGCA	TTTGAGACT	4260
65	GTACCATGCT	GACCATTGCC	CATCGCTGCG	ACACGGTCTC	AGGCTCCGAT	AGGATTATGG	4320
	TGCTGGCCCA	GGGACAGGTG	GTGGAGTTTG	ACACCCCATC	GGTCTCTCTG	TCCAAGCACA	4380
	GTTCOCGATT	CTATGCCATG	TTTGTCTGCT	CAGAGAACAA	GGTGTCTGTC	AAGGGCTGAC	4440
	TCCTCCCTGT	TGACGAAGTC	TCCTTCTCTT	AGAGCATGTC	CATTCCCTGC	CTGGGGCGGG	4500
	CCCTCATCG	CGTCTCTCTA	CCGAAACCTT	GCCTTCTCTG	ATTTTATCTT	TCGCACAGCA	4560
70	GTTCOGGATT	GGCTTGTGTG	TTTCACTTTT	AGGGAGATGC	ATATTTTGAT	TATTGTATTT	4620
	ATTCCATATT	CATGTAAACA	AAATTTAGTT	TTTGTCTTTA	ATTGCACTCT	AAAAGGTTCA	4680
	GGGAACCGTT	ATTATAATTG	TATCAGAGGC	CTATAATGAA	GCTTTATACG	TGTAGCTATA	4740
	TCATATATA	ATTCTGTACA	TAGCCTATAT	TTACAGTGAA	AATGTAAGCT	GTTTATTTTA	4800
	TATTAAATA	AGCAGTGTGC	TAATAACAGT	GCATATTCTT	TTCTATCATT	TTTGTACAGT	4860
75	TTGCTGTACT	AGAGATCTGG	TTTTGCTATT	AGACTGTAGG	AAGAGTAGCA	TTTCTTCTT	4920
	CTCTAGCTGG	TGGTTTACAG	GTGCCAGGTT	TTCTGGGTGT	CCAAAGGAAG	ACGTGTGGCA	4980
	ATAGTGGGCC	CTCCGACAGC	CCCTCTGCCC	GCCTCCCCAC	AGCGCTCCCA	GGGGTGGCTG	5040
	GAGACGGGTG	GGGGCTGGA	GACCATGCGA	AGCGCGGTGA	GTCTCAGGG	CTCCTGCCTT	5100
	CTGTCTGTGT	GTCACTTACT	GTCTCTGTCA	GGAGAGCAGC	GGGGCGAAGC	CCAGGCCCTC	5160
80	TTTCACTCCC	TCCATCAAGA	ATGGGGATCA	CAGAGACATT	CCTCCGAGCC	GGGGAGTTTC	5220
	TTTCTGCGCT	TCTTCTTTT	GCTGTGTGTT	CTAAACAAGA	ATCAGTCTAT	CCACAGAGAG	5280
	TCCCACTGCC	TCAGGTTCTC	ATGGCTGGCC	ACTGCACAGA	GCTCTCCAGC	TCCAAGACT	5340
	GTGGTTTCCA	AGCCCTGGAG	CCAACCTGCT	CTTTTGGAG	TGGCACTTTT	TCAATTCGCT	5400
	ATTCCACAC	CTCCACAGTT	CAGTGGCAGG	GCTCAGGATT	TGTTGGGTCT	GTCTTCTCTT	5460
	CTCACGCGAG	TGCTGCGACA	GTCTCTCTCT	CTCTCTCCCC	TCAAAGCTCG	CAACTTAACT	5520
85	CAGCTCTGCG	TAATCAGTGT	CTCACACTGG	CGTAGAAGTT	TTTGTACTGT	AAAGAGACCT	5580
	ACCTCAGGTT	GCTGGTGTGT	GTGTGGTTTG	GTGTGTTCCC	GCAACCCCC	TTTGTGCTGT	5640
	GGGGCTGGTA	GCTCAGGTGG	GCGTGGTCAC	TGCTGTCTAT	AGTTGAATGG	TCAGCGTTGC	5700



**PCT/US02/12476**

5

[illegible]

40

1	11	21	31	41	51	
AGCAGTTGCA	CAACTTCCAG	CAACTTTCFC	AGCGGGCTAC	TAATGAGCTG	AAAGCCAGGA	60
ACATTCGAGG	AGAAAGAGAG	GCTTCCAGCC	CTCTCTCCCTT	ACACCGTGGAA	CTCCAGACAG	120
CCCCACCCCC	ACCTCAGAA	CACCTTTAAGA	TATTTCTCTT	ATTGCTTGGC	CCGACGACACC	180
ATGGCTCCCT	TGGAGAGAAA	CTTGCTAAAG	ACTCGGCGATA	AAAACAGATC	TCCAATCTAAA	240
GATCATGTATT	CAGAAGAGAA	GGAAATTGTG	GTTTGGGTAT	GCCAAAGAGA	GAACTCTGTC	300
TGTGGGCTGA	CTAAAGCGAC	CACCTCTGCT	GATGTGATCC	AGGCTTTGCT	TGAGGAACAT	360
GAGGCTACGT	TGTGAGAGAA	ACGATTCTCT	CTGGGGGACG	CCAGTGATTA	TGATCATATA	420
GAGAAGTGGG	GAGGCTCGGA	AAGGGTCTTT	CTCCCTCAAC	TGAAATCTCT	CAGGCTTTGG	480
AAAGCGTGGG	GAGATGAGCA	GCCCAATATG	CAATTTTCTT	TGTTTAAAGC	AGATGCTTTT	540
CTTCCAGTTC	CTTTGTGGCG	GACAGCTGAA	GCCAAATTAG	TGCAAAACAT	AGAAAAATTG	600
TGGGAGCTCA	GCCCGACAAA	TCATCAGAG	ACTTTACAC	CAGATAAAC	AAAAAGAATA	660
GTCAGGAAA	TCTTCGGGAA	ACTGGCTAAA	ATTAGCAGC	ACAAGTTCT	TCATGTCTGA	720
GATAATATGG	AGACATAGT	TCATCTGATC	ATTTCGACG	ACCATACAT	TCTCAGCAA	780
GTCAGAGAG	TGAAAGATGT	GGATCTGAA	ATTGAAAGT	GTGAAGCTAA	TGTCATCTT	840
GATCGAGTAT	AAAAATGATG	AGAAAACAT	GTTCAGGAGT	CATATTTAAT	GCCCAAGTTT	900
AGTGAAGTTG	AGCAAAATCT	AGACTTCGAG	GATCGAGAAA	ACCAGACTCT	GGAGGACCTG	960
AGCAAGAGTG	ATGGAATTGA	ACAGCTGGAG	GAACTCAGTA	AATATTACCG	AATACTCAT	1020
GATCAAGCTCT	CTGCTGAAT	AGAAAAGAG	GAAAAGATG	TTGGCATTGA	TATAAATTAA	1080
GATCGCGAAG	GGGAAGCTCG	AAGTGAAGTG	GAAAGCTCTA	ATTTGAAGAG	TGTTAAGTGT	1140
GATTTGGAGA	AAAGCATPSA	AGCTGTGTTT	AAAATTCACT	CTCATTTGAG	TGGTACCTG	1200
AAAGGAGATTA	ATACAGTGA	CTCATTTGCT	CAGATGAAAG	CAAAAGAATA	TGAATCCTG	1260
GCCAAAGAA	TCAATCTACT	TCACATTAGC	AACAAGATG	GGTGCCAGTT	AAAGGAAAAA	1320
AGGACGAAAG	ATCTCAGCT	TCCGAGTAGC	AATGGGGAGA	TCTCTCCCT	TACTCAAGA	1380
GTATTTAGCA	ATTACACAA	TGACACAGAC	TCGGACAGCT	GTATCAGTCT	TAAACCAAGT	1440
CAGGATCTCC	AAACACAGT	AGGAGATGTG	GTGCTGTTGT	CAACATAGT	CCATCGECT	1500
CTTTCTGACC	TGCTTTCTATG	TTTTAATGTT	TGTTTAAATT	AATAGGAAAC	CTCATTTTAA	1560
ATAATACACT	CAAAAATAATG	TAAATCATAT	TGTAGTATTC	AATAGTTAAT	AAAAACTCGA	1620
GAAATGTGTT	GTTTCTG					

75

1	11	21	31	41	51	
MAPFGRNLLK	TRHKNRSPTK	DMDSEEKEIV	VWVQCEKLV	CSLTKRTTSA	DVIQALLEE	60
EATFGKERPL	LKGPSDYCI	EKWGRSRLV	PPLTRILKX	KAWGDEQPNM	QFVLVKDADF	120
LPVPLWRAT	AKVLQNTK	WELSPANYM	KLQDQKQRI	VRKTFPRKLK	QVTLVSHDR	180
DNMETLVHLI	ISQDHTIHQQ	VKRMKLEDL	LEKCEAKFRL	DRVNDGNGY	VQDVALMPSF	240
SEVSNQLDL	EENQNTLED	SESDGIQLE	ERLKYRYRLI	DKLSAEIEKE	VKSCVIDINE	300
DAEGEAASEL	YSSNLESVK	DLEKSMKAGL	KHSHLSGQI	KEIKYSDSLL	QMKAKSYELL	360
QAFNSLHIS	NKDCQCLKEN	RAKESSEVPS	NGEIPPTQR	VFSNYTNDTD	SDTGISNNHS	420
QDEFTVGVDF	VLLGST					

85

Coding sequence: 317-1123

	1	11	21	31	41	51	
5	AGCATTGAAG	GGGAAGGAAC	TGCGGGTGTG	GTGTGTGTAT	GTGTGTGTGT	ATGTGTGTGC	60
	GGCGGTGTGG	TGCGGTGTGTG	TGCGCGCGCT	AGTGTGTGGA	CAAGGAGGTG	GGGGCAGCTG	120
	AGTTAGAGTC	CCAACCTCTG	GACTCCATTT	GCTATTCTCT	TCTTTCTCCC	CCACACCTAT	180
	CTGGTGGTGG	TAGTGGGCGT	TTATATTTCG	GTTCCTTTTC	ATTCAATTTCT	AAATCTCTTA	240
	AAAATTTTGG	GTGGGGGTA	TTGGGGAAGG	CAGGAAAGGG	AAAAGGAGAG	TAGTAGCTGA	300
10	AGAGCAAGAG	GAGGACATGG	AGATGAAGAA	GAAGATTAACT	CTGGAGTTAA	GGAACAGATC	360
	CCCGGAGGAG	GTGACAGAGT	TAGTCCTTGA	TAATTGCTTG	TGTGTCAATG	GGGAAATTGA	420
	AGGCCTGAAT	GATACCTTTCA	AAGAACTAGA	ATTTCTGAGT	ATGGCTAATG	TGGAACTAAG	480
	TTGCGTGGCC	CGGCTTCCCA	GCTTAAATAA	ACTTCGAAAA	TGGAGCTTA	GTGATAATAT	540
	AATTTCTGGA	GGCTTGAAG	TGCTGGCAGA	GAAATGTCCA	AATCTTACCT	ACCTCAATCT	600
15	GAGTGGAAAC	AAAATAAAG	ATCTCAGTAC	AGTAGAAGCT	CTGCAAAATC	TTAAAAATTT	660
	GAAAAGTCTT	GACCTGTTTA	ACTGTGAGAT	CACAAACCTG	GAAGATTATA	GAGAAAGTAT	720
	TTTTGAACCT	CTGCAGCAAA	TCACATACCT	AGATGGATTT	GATCAGGAGG	ATAATGAAGC	780
	GCCGCTGACT	GAAGAGGAGG	ATGATGAGGA	TGGAGATGAA	GATGATGAAG	AGGAAGAGGA	840
20	AAATGAAGCT	GGTCCACCGG	AAGGATATGA	GGAAGAGGAG	GAGGAAGAGG	AAGAGGAGGA	900
	TGAGGATGAG	GATGAAGATG	AAGATGAAGC	AGGTTCAAGG	TTGGGAGAGG	GAGAAGAGGA	960
	AGTGGGCTCT	TCATCTTAA	TGAAAGAAGA	AATTCAGGAT	GAAGAAGATG	ATGATGACTA	1020
	TGTTGAAGAA	GGGAAGAAG	AGGAAGAAGA	GGAAGAAGGA	GCTCTTGGAG	GGGAGAAGAG	1080
	GAAACGAGAT	GCTGAAGACG	ATGGAGAGGA	AGAAGATGAC	TAGATCAATC	TAAGACCAGA	1140
25	TTCTCTAATG	TTTCTGGGTG	TGCAATAGAG	TGATCACATC	TTTGTCTCTT	CATGTACGAT	1200
	AGCTATCCCT	ACAGAAGATA	ATGTGTAAGT	TTTTATAGGA	AAAGTGTGGT	TTTACTATTT	1260
	TTGCTTATTC	ATTTCAAAAT	AGAACTAGTC	TGTTAATGAT	CATATTGTAT	GTAGAGAAAA	1320
	ATTTTCTATT	ACTCCCATTT	TGGAATTTCC	TAGCAATTTA	TTTAGACTTA	ATTTTAAAAA	1380
	TTCAAGCTTA	CTGTATTAGT	CAITTTTAGC	CCATAATTAA	AACATGATCA	CTTTTAAACA	1440
30	GGTGTAGTAT	GGTGCAATTC	ATTCCTTATT	TATAGATTAA	CTGAAATTTAC	AGTTTGGCTAT	1500
	AAATAAAAAA	GACATAGTCT	TCTTGAGTGG	TAAGTTGGTT	ATTTTCTTAG	AGGTGATCCA	1560
	GGAACTTTTA	ATTTCAAGGC	AGTTACCTTT	TTTTTTTTTT	TTTTTTTTTG	ACTAAGAGTG	1620
	TTTGTGTGCT	TTTTTGTGAC	AAGTAACCTG	GAAATAGAAA	GCAGAAATAGT	AAAGTTCTTA	1680
	TTGAGCAACA	TAGTTCTATG	ATTTTGTGGA	GGTCTTATTC	AGTAATATGG	TTTATGGATT	1740
35	TAGTGTGTAC	TGATAAGATT	TTATTTTGA	AGGAAAAATT	GCTTATACTA	AGTCCAGAGA	1800
	CATGCAGGTG	AGCCCTTTTG	TCAGGCTGCA	AATCATGACA	TGCCGATGGT	TGTTTATTTT	1860
	GTTTTATAGT	GTGCAATCTT	TTTCTCTTA	GCAATTCCTT	TATGATCACC	TTCCCTTCTT	1920
	GTTCCTCTCC	CTCCCGCTCT	CTCAAAAGGA	ACTTGGGAAA	CTTGTGAATC	CCAGGAAAAAC	1980
	CTTTAGTCTT	ATACCTCAAC	TACGTTTCAG	TCCGTCTGCG	GTTTTAAATA	AGTGAAGTAG	2040
40	AAGAAATTGA	GTATTTTCTG	ACATAAGAAT	ATATTATCAA	TACAGTTTAA	TGCAGTAAGC	2100
	TCTCCTTACC	ATAAATGTTT	CTTGGTTGAC	AACATCTAAG	ACAATATTAG	TGGGATGAAG	2160
	AAAGAAAGAG	AGGGGTGCTT	TTGGAAGCAG	TGTTAGTGTG	CCTCAAAAGT	CGGAACAATT	2220
	GCCTGTGTAT	ATATTAATAA	GACATTAAAG	TCAAATTTTA	ATGTTGGCCT	CTCAAATGAT	2280
	TTGATACCA	CTCTGCAAG	TATTTCTAAC	CTTTAATTC	CAGTTTAAAA	ACAGATATAA	2340
45	TAATAGCAAT	TAATGGGAAT	ATACTAGCCA	GCTGGAAAAG	TATTTGAAAC	TAAATTGACA	2400
	TTAAATTTAA	GATTTGTTTT	CAAGTGGATG	TCCATTAAAA	GTAGAAAAAT	ATTTGGGATA	2460
	AGTGAGGTG	TGTTTCTCTA	CATGGCTACT	AAATAAAATA	TAATGAGTAT	ACAAGTATAT	2520
	CTCCTCTTTT	GCTATGGAGG	CTCCATGTTT	AAGGCAATGG	CTTTTAAAT	CTTGGCTATC	2580
	TAAATTTTTT	TCCCTTTGTT	TTGAATATTT	GTAAGTTTTT	AAGAAGTTAG	TGTGAGCAAA	2640
50	TTAATGAAG	TTATGCTTCT	ATACTGGGAC	ATATTTAAAT	ACTGAGTATA	GTACTGCTGC	2700
	TACTGCTTCT	ACAATGTAAA	ATGATGACT	TGGTGTTTTA	AAGTAAAAAT	TATGATGTTA	2760
	CTTGTGGAGA	AACTAAAAAT	GTGTACAAC	TGACCGAAAG	AAAACCCCTG	GGGATAAGTT	2820
	TAGTGAGGGG	ATTGGAATCC	CCAAAAGAGT	AACATTTTTC	TTCTGCTTTT	AAAACTGAA	2880
	ATTCCCTGTT	CTAGTTCTTA	ACAATTCTCA	TTACATACTA	TGCCAGATTA	CAAAATACTT	2940
55	ATTTTAAAAA	TGAATCTAT	ATATTGACTT	TCTTATCAAT	CATCTTACTG	TGCAATCAAA	3000
	ATTAGAGTAC	TTTGGTTTGA	AAACAACACT	TAGAGCTCC	AGATAACTTT	TAAGACTTAT	3060
	TTAGCTTTGT	GGGTGGTATT	TTCATGCAAA	TAAGTAAGGG	TGGGTTTTAT	ATTTTGTAGA	3120
	AGTTTTCGAT	CCATTTTAA	TGCTTCTTGT	ATGGCAGTAT	GTATATATTG	TGTTAAGTTC	3180
	CTCAAGAAATC	TCCTTAAAAA	CTTTGAAGTT	AATACCTTTG	TGCAACTGTG	TTTTGAATAA	3240
60	AGCCATGACA	GTGTTAAAAA	CAAAC				

Seq ID NO: 323 Protein sequence:  
Protein Accession #: NP\_112182.1

	1	11	21	31	41	51	
65	MEMKKKINLE	LRNRSPEEVT	ELVLDNCLCV	NGBIEGLNDT	FKELRFLSMA	NVELSSLARL	60
	PSLNKLRLKE	LSDNIISGGL	EVLAERCPNL	TYLNLSGNKI	KDLSTVBALQ	NLQNLKSLDL	120
	FNCEITNLED	YRESIFELLQ	QITYLDGFDQ	EDNEAPDSEE	EDDEDGDEDD	EEEEENEAGP	180
	PEGYEEEEEE	EBEBEDEDEE	DEDBAGSELG	EGEEVGLSY	LMKEEIQDEE	DDDDYVEEGR	240
70	EEEEEEGGL	RGRKRRDAE	DDGEEEDD				

Seq ID NO: 324 DNA sequence  
Nucleic Acid Accession #: NM\_003812  
Coding sequence: 224..2722

	1	11	21	31	41	51	
75	TCCTCTGCGT	CCCGCCCCGG	GAGTGGCTGC	GAGGCTAGGC	GAGCCGGGAA	AGGGGGCGCC	60
	CCCCAGCCCC	GAGCCCCCGG	CCCCGTGCCC	CGAGCCCCGA	GCCCCCTGCC	CGCGCGCGCA	120
80	CCATGCGCGC	CGAGCGCGGC	TGACCGGCTC	CGCCCGCGGC	CGCCCGCGAG	CTAGCCCGGC	180
	GCTCTGCGCG	GCCACACGGA	GCGCGCGCGC	GGAGCTATGA	GCCATGAAGC	CGCCCGCGAG	240
	CAGCTGCGCG	GAGCGCGCCC	TGGCGGGCTG	CAGCCTTGCC	GGCGCTTCTT	GCGGCCCCCA	300
	ACGCGGCCCC	GCGGCTCGGG	TGCCTGCCAG	CGCCCGCGGC	CGCAGCGCGC	CCTGCGCGCT	360
	CTTCTGCTGC	CTTCTCTCTG	TGCTCTCGCT	CGCGCTCTCG	TCCCGGCCCC	GCGCTCGGGG	420
85	GGCTGCTGCG	CCGAGCGCTC	CGCATTGGAA	TGAACTGCA	GAATAAAATT	TGGGAGTCTT	480
	GCCAGATGAA	GACAATACAT	TGCAACAGAA	TAGCAGCAGT	AATATCAGTT	ACAGCAATGC	540
	AATGCAGAAA	GAAATCACAC	TGCCCTCAAG	ACTCATATAT	TACATCAACC	AAGACTCGGA	600

	AAGCCCTTAT	CACGTTCTTG	ACACAAAGSC	AAGACACCAG	CAAAAACATA	ATAAGGCTGT	660
	CCATCTGGCC	CAGGCAAGCT	TCCAGATTGA	AGCCTTCGGC	TCCAAATTCA	TTCTTGACCT	720
	CATACATGAAC	AATGGTTTGT	TGCTTCTGA	TTATGTGGAG	ATTCATACG	AAAAATGGAA	780
5	ACCACAGTAC	TCTAAGGGTG	GAGAGCACTG	TTACTACCAT	GGAAGCATCA	GAGGCGTCAA	840
	AGACTCCAAG	TGGGCTCTGT	CAACCTGCAA	TGGACTTCAT	GGCATGTTTG	AAGATGATAC	900
	CTTGTGTAT	ATGATAGAGC	CACAGAGCT	GTTTCATGAT	GAGAAAAGCA	CAGGTGACCC	960
	ACATATAATC	CAGAAACCTT	TGGCAGGACA	GTATTCTAAG	CAAAATGAAGA	ATCTCATTAT	1020
	GGAAAGAGGT	GACCAATGGC	CCTTCTCTC	TGAATTACAG	TGGTTGAAA	GAAGGAAGAG	1080
10	AGCAGTGAAT	CCATCACGTG	GTATATTGGA	AGAAATGAAA	TATTTGGAAC	TTATGATTGT	1140
	TAATGATCAC	AAAAAGTATA	AGAAGCATCG	CTCTTCTCAT	GCACATACCA	ACAACCTTTC	1200
	AAAGTCCGTG	GTCAACCTTG	TGGATTCTAT	TTACAAGGAG	CAGCTCAACA	CCAGGTTTGT	1260
	CCTGTGGCT	GTAGAGACCT	GGACTGAGAA	GGATCAGATT	GACATCAACA	CCAACCTCTG	1320
	GCAGATGCTC	CATGAGTTCT	CAAAATACCG	GCAGCGCATT	AAGCAGCATG	CTGATGCTGT	1380
15	GCACCTCATC	TCCGGGGTGA	CATTTCCTA	TAAGAGAAGC	AGTCTGAGTT	ACTTTGGAGG	1440
	TGCTGTCTCT	CGCACAAGAG	GAGTTGGTGT	GAATGAGTAT	GGTCTTCCAA	TGGCAGTGGC	1500
	ACAAATGTTA	TGCAGAGGCC	TGGCTCAAAA	CCTTGGAAATC	CAATGGGAAC	CTTCTAGCAG	1560
	AAAGCCAAAA	TGTGACTGCA	CAGAATCTCT	GGGTGGCTGC	ATCATGGAGG	AAACAGGGGT	1620
	TGCCATTCT	CGAAATTTT	CAAAAGTGAG	CATTTTGGAG	TATAGAGACT	TTTTACAGAG	1680
20	AGGAGGTGGA	GCCTGCCTTT	TCAACAGGCC	AACAAGCTA	TTTGAGCCCA	CGGAATGTGG	1740
	AAATGATATC	GTGGAAGCTG	GGGAGGAGTG	TGATTGTGGT	TTTCATGTGG	AATGCTATGG	1800
	ATTATGCTGT	AAGAAATGTG	CCCTCTCCAA	CGGGGCTCAG	TGCAGCGACG	GGCCCTGCTG	1860
	TAACATATCC	TCATGTCTTT	TTACGCCACG	AGGGTATGAA	TGCCGGGATG	CTGTGAACGA	1920
	GTGTGATATT	CTAGAAATAT	GTACTGGAGA	CTCTGGTCAG	TGCCCAACAA	ATCTTCATAA	1980
25	GCAAGACGGA	TATGCATGCA	ATCAAAATCA	GGGCGCTGCG	TACATGGCGG	AGTCAAGAC	2040
	CAGAGACAAC	CAGTGTCTAGT	ACATCTGGGG	AACAAAGGCT	GCAGGGTCTG	ACAAGTTCTG	2100
	CTATGAAAAG	GCAATACAG	AAGGCACTGA	GAAGGGAAGC	TGCCGGAAGG	ATGGAGACCG	2160
	GTGGATTGAG	TGCAGCAAC	ATGATGTGTT	CTGTGGATTC	TTACTCTGTA	CCAATCTTAC	2220
	TGAGACTCCA	CGTATTGGCT	AACCTTCAGG	TGAGATCATT	CCAACCTTCT	TCTACCATCA	2280
30	AGGCGGGTGT	ATTGACTGCA	GTGGTGGCCA	TGTAGTTTAA	GATGATGATA	CGGATGTGGG	2340
	CTATGTAGAA	GATGGAACGC	CATGTGGCCC	GTCTATGATG	TGTTTAGATC	GGAAGTGCCT	2400
	ACAAATTCAA	GCCTTAAATA	TGAGCAGCTG	TCCACTCGAT	TCCAAGGGTA	AAGTCTGTTC	2460
	GGGCGCATGG	GTGTCTAGTA	ATGAAGCCAC	CTGCATTGTT	GATTTTCACCT	GGGCGAGGAC	2520
	AGATTGCACT	TGCTGGGATC	CAGTTAGGAA	CCTTCAACCC	CCCAAGGATG	AAGGACCCAA	2580
35	GGGTCTTAGT	GCCACCAATC	TCATAATAGG	CTCCATCGCT	GGTGCCATCC	TGGTAGCAGC	2640
	TATTTGCTCT	GGGGSCACAG	GCTGGGGATT	TAAAAATGTC	AAGAAGAGAA	GTTTCGATCC	2700
	TACTCAGCAA	GGCCCATCT	GAATCAGCTG	CGCTGGATGG	ACACCGCCTT	GCACGTGTGG	2760
	ATTCTGGGTA	TGACATACCT	GCAGCAGTGT	TACTGGAAGT	ATTAAGTTTG	TAAACAAAAC	2820
	CTTTGGGTGG	TAATGACTAC	GGAGCTAAAG	TTGGGGTGAC	AAGGATGGGG	TAAAGAAAAA	2880
40	CTGTCTCTTT	TGGAATTAAT	GTCAAGAAGC	ACCTTTCACC	ACCTGTCAGT	AAACGGGGGA	2940
	GGGGGCAAAA	GACCATGCTA	TAAAAAGAAC	TGTTCCAGAA	TCTTTTITTT	TCCCTAATGG	3000
	ACGAAGGAAC	AACACACACA	CAAAAATTAA	ATGCAATAAA	GGAATCATTAA	AAAA	

Seq ID NO: 325 Protein sequence:  
Protein Accession #: NP\_003803

45	1	11	21	31	41	51	
	MKPPGSSSRQ	PPLAGCSLAG	ASCGPQRGPA	GSVPASAPAR	TPPCRLLLV	LLLPLAASS	60
50	RPRAGAAAP	SAPHWNETAE	KNLGLVADED	NTLQNNSSSN	ISYSNAMQKE	ITLPSRLIYY	120
	INQDSSEPYH	VLDTKARHQQ	KHNKAHLAQ	ASFQIEAPGS	KFILDILN	GLLSSDYVEI	180
	HYENGKPYQS	KGGHECYHSG	SIRGVKDSKV	ALSTCNGLHG	MFEDDTFVYM	IEPLELVEDE	240
	KSTGRPHIIQ	KTLAGQYSKQ	MKNLTMERGD	QNPFLSELQW	LKRRKRAVNP	SRGIFEEMKY	300
	LELMIVNDHX	TYKKHRSSEA	HTNPNFAKSVV	MLVDSIYKEQ	LNRVRLVAV	ETWTEKDID	360
55	ITTNPVQMLH	EPFSKYRQIK	QHADAHLIS	RVTPHYKRSS	LSYPGGVCSR	TRGVGVNEYG	420
	LPMAVAQVLS	QSLAQNLGIC	WEPSSRKPKC	DCTESWGGC	MEETGVSHSR	KFSKCSILEY	480
	RDFLRGGGA	CLFNRPRTKLF	EPTGCGNGYV	EAGEECDCGF	HVECYGLCK	KCSLSNGAHC	540
	SDGPCNNNTS	CLFPQPRGYEC	RDVNECDIT	EYCTGDSGQC	PPNLHKQDGY	ACNQNGRCY	600
	NGECKTRDNQ	CQYINGTRAA	GSDKFCYEKL	NTEGTERKNC	GKDGDRWIC	SKHDVFCGFL	660
60	LCTNLTRAPR	IQGLQGETIP	TSFYHQGRVI	DCSGAHVLLD	DDTDVGVYED	GTCPGSPMMC	720
	LDRKCLQIQ	LMNMSCPLDS	KGKVCSEHGV	CSNEATCID	FTWAGTDCSI	RDPVRNLHPP	780
	KDBGPKGPSA	TNLIGSIAG	AILVAIVLG	GTGWGFRNVK	KRRFDPTQGG	PI	

Seq ID NO: 326 DNA sequence  
Nucleic Acid Accession #: AK074418.1  
Coding sequence: 244-1515

65	1	11	21	31	41	51	
	CTTTCTCCAA	GAGCGCCGCG	CATGCTCTCC	TCCTCTGCCA	GTCTCTCTCA	CCACTCTCTA	60
70	ACCTGAGAGC	CTGTGGAACC	TGCCCGTCTC	CCCTCTCTCA	TCAGACACAC	CTGCCTAGGA	120
	AACAGATGGA	AAAAGTGAGG	GACCGGTGAG	TGACTTCTGT	CTAAAGTTTA	TACCAGATGC	180
	AAATGACAGA	GCTGGAGTTC	TGCTGTGCCT	GGAAAGGACC	TGGAAGTCT	TCTAAGGAGA	240
	GTCTGSGGT	ATTACACAGG	GCCTTCAGTG	GAGACCTCCA	TCATCAGTT	CAAAGACGAG	300
75	GACTTTACCA	CCTTGGCGGA	TCACTGCCTG	AGCATGGGCC	GGACGTTTAA	GGATGAGACA	360
	TTCCCGCAG	CAGATTCTTC	CATAGGCCAG	AAGCTGCTCC	AGGAAAAACG	CCTCTCCAAT	420
	GTGATATGGA	AGCGGCCACA	GGATCTACCA	GGGGGTCTCT	CTCACTTCAT	CCTGGATGAT	480
	ATAAGCAGAT	TTGACATCCA	ACAAGGAGGC	GCAGCTGACT	GCTGGTCTCT	GGCAGCACTG	540
	GGATCTCTGA	CTCAGAACC	ACAGTACAGG	CAGAAGATCC	TGATGGTCCA	AAGCTTTTCA	600
80	CACGAGTATG	CTGGCATTTT	CGGTTCCGG	TTCTGGCAAT	GTGGCCAGTG	GTTGGAAGTG	660
	GTGATTGATG	ACCGCTTACC	TGTTCCAGGA	GATAAATGCC	TCTTTGTGCG	TCTCTGCCAC	720
	CAAAACCAAG	AGTCTTGCC	CTGCCGTCTG	GAGAAGGCCT	ATGCCAAGCT	GCTCGGATCC	780
	TATTCGATC	TGCACTATGG	CTTCTCGAG	GATGCCCTGG	TGGAACCTCAC	AGGAGCGGTG	840
	ATCACCAACA	TCCATCTGCA	CTCTTCCCT	GTGGACCTGG	TGAAGGCAGT	GAAGACAGCG	900
85	ACCAAGGCAG	GCTCCCTGAT	AACCTGTGCC	ACTCCAAGTG	GGCCAACAGA	TACAGCACAG	960
	GCGATGGAGA	ATGGGCTGGT	GAGTCTCCAT	GCCTACACTG	TGACTGGGCG	TGAGCAGATT	1020
	CAATACCGAA	GGGGCTGGGA	AGAAATTATC	TCCTGTGGGA	ACCCCTGGGG	CTGGGCGGAG	1080
	ACGAATGGA	GAGGCGCTG	GAGTGATGGG	TCTCAGGAGT	GGGAGGAAAC	CTGTGATCCG	1140

	CGGAAAGCC	AGCTACATAA	GAACCGGAA	GATGGCGAGT	TTTGGATGTC	GTGTCAAGAT	1200
	TTCCAACAGA	AATTCATCGC	CATGTTTATA	TGTAGCGAAA	TTCCAATTAC	CCTGGACCAT	1260
	GGAAACACAC	TCCACGAAGG	ATGGTCCCAA	ATAATGTTTA	GGAAAGCAAGT	GATTCTAGGA	1320
5	AACACTGCGAG	GAGGACCTCG	GAATGATGCT	CAATTCAACT	TCTCTGTGCA	AGAGCCAATG	1380
	GAAGGCACCA	ATGTTGTCTG	GTGGCTCACA	GTGTCTGTCA	CACCATCAAA	TTTGAAGCA	1440
	GAAGATGCAA	AATTTCCACT	CGATTTCCAA	GTGATTCTGG	CTGGCTCACA	GAACACTGT	1500
	CCAAAGCTCA	AATAATAAAT	TCCGCCGCAA	CTTCACCATG	ACTTACCATC	TGAGCCCTGG	1560
	GAACATATGT	GTGGTTGCAC	AGACACGGAG	AAAATCAGCG	GAGTTCCTTC	TCCGAATCTT	1620
10	CCTGAAAATG	CCAGACAGTG	ACAGGCACCT	GAGCAGCCAT	TTCAACCTCA	GAATGAAGGG	1680
	AAGCCCTTCA	GAACATGGCT	CCCAACAAAG	CATTTTCAAC	AGATATGCTC	AGCAGGTATG	1740
	GTACCTAGCA	CCCAAGGGCC	TTACGTGGGA	TTGGAGAAAG	GGGACCTGAG	GGAGGGACAG	1800
	CCCTCAGAG	CCCTTACTGG	GATGCAGAGA	GGAGAAGTGA	CTTGATGGAC	TATTTTACCT	1860
	GCCTCTCTTC	CTGGATCGTC	TCCAGAACTG	CTGTGGCTGC	CAAGCTCGGT	AGAGACGTGG	1920
	CGCCCCACCC	AGTCTCATCC	GGGGGACTTC	AAGCTGGAAT	GCAGAGCTTA	GAAGGGGAGG	1980
15	GGATAATTAT	GGGGTGTGAG	GTGCATTGCC	CTCTAAATCT	TTAAACAAGC	AATTGGCAGT	2040
	ACCCCGTGAA	ACCTTTCTCT	CTCCTACTCG	GCCACCTCCC	ACCAACCTGG	CATCGTTCTT	2100
	CCCGGAGAGT	AGCCAGCTTC	AGAAAGCACA	TACAGCATCC	TTGCTGCCAA	ACCACCTATG	2160
	TGCACACAGG	ATTTCCTTAA	TGGCTTAATA	AACCTGTATA	AAGAACCTCT	TGACTTGTCA	2220
20	GAATAAAATA	GCTGCCAGGG	GCTCTGCACA	ATGAGCCTCT	TACCGTTAAA	AAAAAAAATA	2280
	AAAAAAAATA	AAAAAAAATA	AAAAAAAATA				

Seq ID NO: 327 Protein sequence:

Protein Accession #: BAB85075.1

25	1	11	21	31	41	51	
	MAYYQEPSVE	TSIIKFKQDQ	FTTLRDHCLS	MGRTFKDETF	PAADSSIGQK	LLQEKRLSNV	60
	IWKRPQDLPG	GPPHFILDDI	SRFDIQQGGG	ADCNFLAALG	SLTONPQYRQ	KILMVQSPSH	120
30	QYAGIFRFRF	WQCGQWVEVV	IDDRLPVQGD	KCLFVRPRHQ	NQEPNWCLE	KAYAKLIGSY	180
	SDLHYGFLED	ALVDLTGTVI	TNIHLHSSPV	DLVKAVKTAT	KAGSLITCAT	PSGPIDTAQA	240
	MENGLVSLHA	YTVTGAEIQI	YRRGWEEIIS	LWNPWNGWET	EWNRGRWSDGS	QEWEECTDPR	300
	KSQLHKHRED	GEFWMSCQDF	QOKFIAMPIC	SEIPITLDHG	NTLHEGWSQI	MFRKQVILGN	360
	TAGGPRNDAQ	FNFSVQEPME	GTNVVVCVTV	AVTPSNLKAE	DAKFPLDFQV	ILAGSQKHCP	420
35	KLK						

Seq ID NO: 328 DNA sequence

Nucleic Acid Accession #: BC017490.1

Coding sequence: 74-2788

40	1	11	21	31	41	51	
	GTGGGTCACG	TGAACCACTT	TTGGGCGGAA	ACCTGGTGTG	TGCTGTAGTG	GCGGAGAGGA	60
45	TGCTGGTACT	GCTATGGCGG	AATCATCGGA	ATCCTTCACC	ATGGCATCCA	GCCCGGCCCA	120
	GGGTGGGCGA	GGCAATGATC	CTCTACCTTC	CAGCCCTGGC	CGAAGCTCCA	GGCGTACTGA	180
	TGCGCTCACC	TCCAGCCCTG	GCGGTGACCT	TCCACCATTT	GAGGATGAGT	CCGAGGGGCT	240
	CCTAGGCACA	GAGGGGCCCC	TGGAGGAAGA	AGAGGATGGA	GAGGAGCTCA	TTGGAGATGG	300
	CATGGAAGGG	GACTACCCGG	CCATCCGAGA	GCTGGAGCGC	TATGAGGCGG	AGGGAGCTGGC	360
50	TCTGGATGAT	GAGGACGTAG	AGGAGCTGAC	GGCCAGTCAG	AGGGAGGCGG	CAGAGCGGGC	420
	CATGCGGCAG	CGTGACCGGG	AGGCTGGCGG	GGGCGTGGGC	CGCATGCGCC	GTGGGCTCCT	480
	GTATGACAGC	GATGAGGAGG	ACGAGGAGCG	CCCTGCCCGC	AAGCGCGCGC	AGGTGGAGCG	540
	GGCCACGGAG	GACGGCGAGG	AGGACGAGGA	GATGATCGAG	AGCATCGAGA	ACCTGGAGGA	600
	TCTCAAAGGC	CACCTCTGTC	GCGAGTGGGT	GAGCATGGCG	GGCCCCCGGC	TGGAGATCCA	660
55	CCACCGCTTC	AAGAACTTCC	TGGGCACTCA	CGTCGACAGC	CACGGCCACA	ACGTCTTCAA	720
	GGAGCGCATC	AGCGACATGT	GCAAGAGAGG	CCGTGAGAGC	CTGGTGGTGA	ACTATGAGGA	780
	CTTGGCAGCC	AGGGAGCAAG	TGCTGGGCTA	CTTCTGCTCT	GAGGCACCGG	CGGAGCTGCT	840
	CGAGATCTTT	GATGAGGCTG	CCCTGGAGGT	GGTACTGGCC	ATGTACCCCA	AGTACGACCG	900
	CATCACCAC	CACATCCATG	TCCGCATCTC	CCACCTGCCT	CTGGTGGAGG	AGCTGCGCTC	960
60	GCTGAGGCG	CTGCATCTGA	ACCAGCTGAT	CCGACCCAGT	GGGGTGGTGA	CCAGCTGCAC	1020
	TGGGCTCTCT	CGCCAGCTCA	GCTAGGTCAA	GTACAACTGC	AACAAGTGCA	ATTTCTCTCT	1080
	GGGTCTTTTC	TGCCAGTCCC	AGAACCAGGA	GGTGAAACCA	GGCTCTCTGC	CTGAGTGCCA	1140
	GTCGGCGCGC	CCCTTTGAGG	TCAACATGGA	GGAGACCATC	TATCAGAACT	ACCAGCGTAT	1200
	CGAATCCAG	GAGAGTCCAG	GCAAGTGGC	GGCTGGCGCG	CTGCCCGGCT	CCAAGGACGC	1260
65	CATTCTCTCT	GCAGATCTGG	TGGACAGCTG	CAAGCCAGGA	GACGAGATAG	AGCTGACTGG	1320
	CATCTATCAC	AACAACTATG	ATGGCTCCCT	CAACACTGCC	AATGGCTTCC	CTGTCTTTGC	1380
	CACGTGATC	CTAGCCAACC	ACGTGGCCAA	GAAGGACAA	AAGGTGCTGT	TAGGGGAAC	1440
	GACCGATGAA	GATGTGAAGA	TGATCACTAG	CCTCTCCAAG	GATCAGCAGA	TGGGAGAGAA	1500
	GATCTTTGCC	AGCATTGCTC	CTTCCATCTA	TGGTCAATGA	GACATCAAGA	GAGGCCCTGG	1560
70	TCTGGCCCTG	TTCGGAGGGG	AGCCCCAAAA	CCAGGTGGC	AAGCACAAGG	TACGTGTGTA	1620
	TATCAACGTG	CTCTTGTGCG	GAGACCCCTG	CACAGCGAAG	TCGCAATTTC	TCAAGTATAT	1680
	TGAGAAAGTG	TCCAGCCGAG	CCATCTTCAC	CATGGCCAG	GGGGCGTCGG	CTGTGGGCCT	1740
	CACGGCGTAT	GTCCAGCGGC	ACCTGTTCAG	CAGGGAGTGG	ACCTTGGAGG	CTGGGGCCCT	1800
	GGTTCTGGCT	GACCGAGGAG	TGTGTCTCAT	TGATGAATTT	GACAAGATGA	ATGACCCAGGA	1860
75	CAGAACCAGC	ATCCATGAGG	CCATGGAGCA	ACAGAGCATC	TCCATCTCGA	AGGCTGGCAT	1920
	CGTCACTCCC	CTGCAGGCTC	GCTGCAAGGT	CATTGTGCGC	GCCAAACCCA	TAGGAGGGCG	1980
	CTACGACCCC	TGCTGACTTT	TCTCTGAGAA	CGTGGACCTC	ACAGAGCCCA	TCACTCTCAG	2040
	CTTTGACATC	CTGTGTGTGG	TGAGGGACAC	CGTGGACCCA	GTCCAGGACG	AGATGCTGGC	2100
	CGCTCTGTGT	GTGGGCGAGC	ACGTGACACA	CCACCCGAGC	AACAAGGAGG	AGGAGGGGCT	2160
80	GGCCAAATGG	AGCGCTGCTG	AGCCCGCCAT	GCCCAACACG	TATGGCGTGG	AGCCCTGCCC	2220
	CCAGGAGGTC	CTGAAGAAGT	ACATCATCTA	CGCCAAAGGAG	AGGGTCCACC	CGAAGCTCAA	2280
	CCAGATGGAC	CAGGACACAG	TGGCCAAAGT	GTACAGTGAC	CTGAGGAAAG	AATCTATGCG	2340
	GACAGGCAGC	ATCCCCATTA	CGGTGCGGCA	CATCGAGTCC	ATGATCCGCA	TGGCGGAGGC	2400
	CCAGCGCGCG	ATCCATCTGC	GGGACTATGT	GATCGAAGAC	GACGTCAACA	TGGCCATCCG	2460
85	CGTATGCTGT	GAGAGCTTCA	TAGACACACA	GAAGTTACAG	GTCTGCGCGA	CGATGCGCAA	2520
	GACTTTTGCC	CGTACCTTTT	CATTCCGGCG	TGACAAACAT	GAGCTGTTGC	TCTTCACTACT	2580
	GAAGCAGTTA	GTGGCAGAGC	AGGTGACATA	TCAGCGCAAC	CGCTTTGGGG	CCGACGAGGA	2640
	CACATATGAG	GTCCCTGAGA	AGGACTTGTG	GGATAAGGCT	CGTCAGATCA	ACATCCACAA	2700

WO 02/086443

PCT/US02/12476

CCTCTCTGCA TTTTATGACA GTGAGCTCTT CAGGATGAAC AAGTTCAGCC ACGACCTGAA 2760  
 AAGGAAAAATG ATCCTCGCAGC AGTTCCTGAGG CCCTATGCCA TCCTAAGGA TTCTTGGGA 2820  
 TTCTGGTTTG GGGTGGTTCAG TGCCCTCTGT GCTTTATGGA CACAAAACCA GAGCACTTGA 2880  
 TGAACCTGGG GTACTAGGCT CAGGGCTTAT AGCAGGATGT CTGGCTGCAC CTGGCATGAC 2940  
 TGTCTGTTTC TCCAAGCCTG CTTTGTGCTT CTCACCTTTG GGTGGGATGC CITGCCAGTG 3000  
 TGCTCTACTT GGTGTCTGAA CATCTTGCCA CCTCCGAGTG CTTTGTCTCC ACTCAGTACC 3060  
 TTGGATCAGA GCTGCTGAGT TCAGGATGCC TGCCTGTGCT TTAGGTGTTA GCCTCTTCTA 3120  
 ATGGATGTA GAGAGAGCTG TGCCCTCTTG GCGTGAGTTG CGTATTCAGG CTGCTTTTGC 3180  
 TGCTTTGGC CAGAGAGCTG GTTGAAGATG TTTGTAATCG TTTTCAGTCT CCGCAGGTT 3240  
 TCTGTGCCCC TGTGGTGGAA GAGGSCACGA CAGTGCCAGC GCAGCGTTCT GGGCTCTCTA 3300  
 GTGCGAGGGG TGGGATGTGA GTCATGCGGA TTATCCACTC GCCACAGTTA TCAGCTGCCA 3360  
 TTGCTCCCTG TCTGTTTCCC CACTCTCTTA TTTGTGCTT CCGTTTGGTT TCTGTAGTTT 3420  
 TAATTTTAA TAAAGTTGAA TAAATATAA AAAAAA AAAAAA

Seq ID NO: 329 Protein sequence:  
 Protein Accession #: AAH17490.1

1 11 21 31 41 51  
 MAESSESPTM ASSPAQRRRG NDPLTSSPGR SSRRTDALTS SPGRDLPPFE DSEGLLGTE 60  
 GPLEEEDGE ELIGDGMERD YRAIPELDAY EAEGLALDDE DVEELTASQR EAAERAMRQR 120  
 DREAGRGLGR MRRLGLYSD EDEEREPARK RRQVERATED GREDEEMIES IENLEDLKGH 180  
 SVREWVSMAG PRLEIHRRFK NFLRTHVDSH GHNVPKERIS DMCKENRESL VVNYEDLAAR 240  
 EHVLAFLPE APABELLQIFD EAALEVVLAM YPKYDRITNH IHVRISHLPL VEELRSLRQL 300  
 HLANQLIRTSG VVTSCTGVLF QLSMVKYN CNFVLGPPFC QSQNQEVKPG SCPECCSAGP 360  
 FEVMEETIY QNYQIRIQE SPGKVAAGRL PRSKDAILLA DLVDSCKPGD EIELTGIYHN 420  
 NYDGSINTAN GPPVATVIL ANHVAKDNK VAVGELTDED VKMITSLSKD QQIGEKIFAS 480  
 IAPSIYGHED IKRGLALALP GGEPPKPPGK HKVRGDIINV LCGDPTAKS QPLKIEKVS 540  
 SRAIFTTGGG ASAVGLTAYV QRHPVSEWNT LEAGALVLAD RGVCLIDEFD KMDQDRTSI 600  
 HEAMEQSSIS ISKAGIVTSL QARCTVIAAA NPIGGRYDPS LTFSENVDLT EPIISRFDIL 660  
 CVVRDVTDFV QDEMLARFVV GSHVRHPSN KEEEGLANGS AAEFAMPNTY GVEPLPQEV 720  
 KKYIYAKER VHPKLNQMDQ DKVAKMYSDL RKESMATGSI PITVRHIESM IRMAEAHARI 780  
 HLRDYVIDED VNMAIRVMLE SFIDTQKPSV MRSMRKTFAR YLSFRDNE LLLFILKQLV 840  
 AEQVTYQRNR FGAQDQDTEV PEKDLVDRAR QINIHNSAP YDSELFPMNK FSHDLKRMKI 900  
 LQQF

Seq ID NO: 330 DNA sequence  
 Nucleic Acid Accession #: M17254  
 Coding sequence: 257-1645

1 11 21 31 41 51  
 GTCCGCGGCT GTCCGCGCCC GCGTGTGCCA GCGCGCGTGC CTTGGCGGTG CGCGCGGAGC 60  
 CGGGTGCAC TAACTCCCTC GCGCGCGAGC GCGCGCGTAA CCTCTCGGTT ATTCCAGGAT 120  
 CTTTGAGAGC CCAGAGAAAG CCGTGTGAC CAAAAGCAAG ACAATGACT CACAGAGAAA 180  
 AAAGATGGCA GAACCAAGGG CACTAAAGC CGTCAGGTTT TGAACAGCTG GTAGATGGGC 240  
 TGGCTTACTG AAGGACATGA TTCAGACTGT CCGGACCCA GCAGCTCATA TCAAGGAAGC 300  
 CTTATCAGTT GTGAGTAGAG ACCAGTCGTT GTTTGAGTGT GCCTACGGAA CGCACACCT 360  
 GGCTAAGACA GAGATGACCG CGTCCCTCTC CAGCGACTAT GGACAGACTT CCAAGATGAG 420  
 CCCACGCTC CTTCCAGCAG ATTGGCTGTC TCAACCCCCA GCCAGGTC CAATCAAAAT 480  
 GGAATGTAAC CCTAGCCAGG TGAATGGCTC AAGGAACCTT CTTGATGAAT GCAGTGTGGC 540  
 CAAAGCGGGG AAGATGTTGG GCAGCCCAAG CACCGTTGGG ATGAACACG GCAGCTACAT 600  
 GGAGGAGAGG CACATGCCAC CCCCACCAT GACCACGAAC GAGCGCAGAG TTATCGTGCC 660  
 AGCAGATCTT ACGCTATGGA GTACAGACCA TGTGCGGCGAG TGGCTGGAGT GGGCGGTGAA 720  
 AGAATATGGC CTTCCAGAGG TCAACATCTT GTTATTCAG AATCGATG GGAAGGAAT 780  
 GTCAAGATG ACCAAGGAGG ACTTCCAGAG GCTCACCCCC AGCTACACG CCGACATCCT 840  
 TCTCTCATAT CTCCTACTAC TCAGAGAGAC TCCTCTTCCA CATTTGACTT CAGATGATGT 900  
 TGATAAAGCC TTACAAACT CTCCACGGTT AATGCTATGT AGAAACACAG ATTTACCATA 960  
 TGAGCCCCC AGGAGATCAG CCGTGACCGG TCACGCGCAC CCGACGCCCC AGTCGAAAGC 1020  
 TGCTCAACCA TCTCCTTCCA CAGTGCCCAA AACTGAAGAC CAGCGTCTCT AGTTAGATCC 1080  
 TTATCAGATT CTTGAGACCA CAAGTAGCCG CCTTGCAAT CCAGGCGATG GCCAGATCCA 1140  
 GCTTTGGCAG TTCCTCTGTC AGCTCCTGTC GGACAGCTCC AACTCCAGCT GCATCACCTG 1200  
 GGAAGGCACC AACGGGAGT TCAAGATGAC GGATCCCGAC GAGGTGGCCC GCGCTGGGG 1260  
 AGAGCGGAAG AGCAAAACCA ACATGAACCTA CGATAAGCTC AGCGCGGCC TCGTTACTA 1320  
 CTATGACAAG AACATCATGA CCAAGGTCCA TGGGAAGGCG TACGCTTACA AGTTGAGCTT 1380  
 CCACGGGATC GCCCAGGCC TCCAGCCCCA CCCCCCGGAG TCATCTCTGT ACAAGTACCC 1440  
 CTCAGACCTC CCGTACATGG GCTCCTATCA CGCCACCCA CAGAAGATGA ACTTTGTGGC 1500  
 GCCCCACCTT CCGACCTTCC CCGTGACATC TTCCAGTTT TTTGCTGCC CAAACCCATA 1560  
 CTGGAATTCA CCAACTGGGG GTATATACCC CAACACTAGG CTCGCCACCA GCCATATGCC 1620  
 TTCTCATCTG GGCACCTTACT ACTAAGACC TGGCGGAGGC TTTTCCCATC AGCGTGCAAT 1680  
 CACGAGCCCA TCGCCACAAA CTCTATCGGA GAACATGAAT CAAAAGTGCC TCAAGAGGAA 1740  
 TGAATAAAGC TTTACTGGGG CTGGGGAAGG AAGCGGGGA AGAGATCCAA AGACTCTTGG 1800  
 GAGGAGTTA CTGAAGTCTT ACTACAGAAA TGAGGAGGAT GCTAAAAATG TCAAGGAATAT 1860  
 GGACATATCA TCTGTGACT GACCTTGTA AAGACAGTGT ATGTAGAAGC ATGAAGTCTT 1920  
 AAGGACAAAG TGCCAAAGAA AGTGGTCTTA AGAATGTAT AAACCTTAGA GTAGAGTTTG 1980  
 AATCCCACTA ATGCAAACTG GGATGAAACT AAAGCAATAG AAACAACACA GTTTTGACCT 2040  
 AACATACGCT TTATAATGCC ATTTTAAAGG AACTACCTG TATTTAAAAA TAGTTTCATA 2100  
 TCAAAACAA GAGAAAGAC ACGAGAGAGA CTGTGGCCCA TCAACAGACG TTGATATGCA 2160  
 ACTGCATGGC ATGTGCTGTT TTGGTTGAAA TCAATACAT TCCGTTTGAT GGACAGCTGT 2220  
 CAGCTTTCTC AAACCTGAAA GATGACCCAA AGTTTCCAA TCCTTTACAG TATTACCGG 2280  
 ACTATGAAT AAAAGGTGGG ACTGAGGATG TGTATAGAGT GAGCGTGTGA TTGTAGACAG 2340  
 AGGGGTGAG AGGAGGAGG AAGAGGCGAG GAAGGAGGAG ACCAGGCTGG GAAAGAACT 2400  
 TCTCAAGCAA TGAAGACTGG ACTCAGGACA TTTGGGACT GTGTACATG AGTTATGGAG 2460  
 ACTCGAGGGT TCAATGACGT AGTGTATAC CAAACCAAGT GTTAGGAGAA AGGACACAGC 2520  
 GTAATGGAGA AAGGGAAGTA GTAGAATTCA GAAACAAAA TGCGCATCTC TTTCTTTGTT 2580  
 TGTCAAATGA AATTTTAACT TGAATTTGTC TGATATTAA GAGAAACATT CAGGACCTCA 2640  
 TCATTATGTG GGGGCTTTGT TCTCCACAGG GTCAGGTAA AGATGGCCTT CTTGGCTGCC 2700

ACAATCAGAA ATCAGCAGG CATTTTGGGT AGGCGGCCCTC CAGTTTTCCT TTGAGTCGCG 2760  
 AACGCTGTGC GTTTGTGAGA ATGAAGTATA CAAGTCAATG TTTTCCCCC TTTTATATA 2820  
 ATAATATATAT AACTTATGCA TTATACACT ACGAGTGTAT CTCGCCACG CAAAGACACA 2880  
 CGACAAAGA GACAAATCGAT ATAATGTGGC CTTGAATTTT AACTCTGTAT GCTTAATGTT 2940  
 TACAATATGA AGTTATTAGT TCTTAGAATG CAGAATGTAT GTAATAAAAT AAGCTTGGCC 3000  
 TAGCATGGCA AATCAGATT ATACAGGAGT CTGCATTGTC ACTTTTTTTA GTGACTAAAG 3060  
 TTGCTTAATG AAAACATGTG CTGAATGTTG TGGATTGTTT GTTATAATTT ACTTTGTCCA 3120  
 GGAACCTTGG CAAGGGAGAG CCAAGGAAAT AGGATGTTTG GCACCC

Seq ID NO: 331 Protein sequence  
 Protein Accession #: AAA52398

1 11 21 31 41 51  
 | | | | |  
 MIQTVDPDAA HIKEALSUVS EDQSLFECAY GTPHLAKTEM TASSSSDYGO TSKMSPRVPQ 60  
 QDWLSQPPAR VTIKMECNFS QVNGSRNSFD ECSVAKGGKM VGSPTVGMN YGSYMEEKHM 120  
 PPNMTTNER RVIVPADPTL NSTDHRVQWL ENAVKEYGLP DVNILLFQNI DGKELCKMTK 180  
 DDFQLTSPSY NADILLSHLE YLRETEPLPHL TSDVDVKALQ NSPRLMHARN TDLPEYPPRR 240  
 SAWTGHGHT PQSKAAQSPS STVPKTEDQR PQLDYPYILG PTSRLANPG SGQIQLWQFL 300  
 LELLSDSNS SCITWEGTNG EFKMTDPDEV ARRWGERKSK FNMNYDKLSR ALRYYYDKNI 360  
 MTKVHGKRYA YKFDPHGIAQ ALQPHPPSS LYKYPDLFPY MGSYHHPQK MNFVAPHPA 420  
 LPVTSSSFFA APNFWNSPT GGIYPNTRLP TSHMPSHLGT YY 462

Seq ID NO: 332 DNA sequence  
 Nucleic Acid Accession #: NM\_000020  
 Coding sequence: 283-1794

1 11 21 31 41 51  
 | | | | |  
 AGGAAACGTT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60  
 AGAAACATT TTGCTCCAGC CCCCATCCCA GTCCCGGAGG GCTGCGCGGC CAGCTGCGCC 120  
 GAGCGAGCCC CTCCCAGGCT CAGCCCGGCT CCGGGGCGCG GCCGAGCCCC AGCCCGCGCT 180  
 CCAGCCGCTGG CGGTGCAACT GCGCCGCGCG GGTGGAGGGG AGGTGGCCCC GTCCGCGCA 240  
 AGGCTAGCGC CCGGCCACCC GCAGAGCGGG CCCAGAGGGA CCATGACCTT GGGCTCCCCC 300  
 AGGAAAGGCC TTCTGATGCT GCTGATGGCC TTGGTGACCC AGGGAGACCC TGTGAAGCCG 360  
 TCTCGGGGCC CGCTGGTGAC CTGCACGTGT GAGAGCCAC ATTGCAAGGG GCCTACCTGC 420  
 CGGGGGGCGCT GGTGCACAGT AGTGTCTGGT CCGGAGGAGG GGAGGCACCC CCAGGAACAT 480  
 CGGGGCTGGG GGAACCTGCA CAGGGAGCTC TGCAGGGGCG GCCCAGCGA GTTCGTCAAC 540  
 CACTACTGCT GCGACAGCCA CCTCTGCAAC CACAACGTGT CCTTGGTGCT GGAGGCCACC 600  
 CAACCTCTCT CGGAGCAGCC GGGAAACAGT GGCAGCTGG CCTGATCCT GGGCCCCGTG 660  
 CTGGCCTTGG GGCCTCTGGT GGCCTTGGGT GTCTTGGGCC TGTGGCATGT CCGACGGAGG 720  
 CAGGAGAGGC AGCGTGGGCT GCACAGCGAG CTGGGAGAGT CCAGTCTCAT CCTGAAAGCA 780  
 TCTGAGCAGG GCGACACGAT GTTGGGGGAC CTCCTGGACA GTGACTGCAC CACAGGGAGT 840  
 GGCTCAGGGC TCCCTTCTCT GGTGCAGAGG ACAGTGGCAC GGCAGGTTGC CTGTGTGGAG 900  
 TGTGTGGGAA AAGGCCGCTA TGGCGAAGTG TGGCGGGGCT TGTGGCAGG TGAGAGTGTG 960  
 GCGGTCAAGA TCTTCTCTCT GAGGGATGAA CAGTCTCTGT TCCGGAGAGC TGAGATCTAT 1020  
 AACACAGTAT TGCTCAGACA CGACAACATC CTAGGCTTCA TCGCTCAGA CATGACCTCC 1080  
 CGCAACTCGA GCAACGCTAT GTGGCTCATC ACGCACTACC ACGAGCACGG CTCCTCTCAC 1140  
 GACTTCTTGC AGAGACAGAC GCTGGAGCCC CATCTGGCTC TGAGGCTAGC TGTGTCCGCG 1200  
 GCATGCGGCC TGCGCACCT GCACGTGGAG ATCTTCGTA CACAGGCAA ACCAGCCATT 1260  
 GCGCACGGG AGTCTCAAGG CCGCAATGTG CTGGTCAAGA GCAACCTGCA GTGTGTGATC 1320  
 GCGGACCTGG GCTCGGCTGT GATGCACTCA CAGGGCAGCG ATTAAGTGA CATCGGCAAC 1380  
 AACCCGAGAG TGGGCAACAA GCGGTACATG GCACCCGAGG TGCTGGACGA GCAGATCCCG 1440  
 ACGGACTGCT TTGAGTCTCA CAGTGGACT GACATCTGGG CCTTGGCCT GGTGTGTGG 1500  
 GAGATTGGCC GCGGACCAT OGTGAATGGC ATCGTGGAGG ACTATAGACC ACCCTTCTAT 1560  
 GATGTGTGTC CCAATGACCC CAGCTTTGAG GACATGAAGA AGGTGTGTG TGTGGATCAG 1620  
 CAGACCCCA CATCTCTTAA CCGGCTGGCT GCAGACCCGG TCTCTCAGG CCTAGCTCAG 1680  
 ATGATGCGGG AGTGTCTGTA CCAAAACCCC TCTGCCGAC TCACCGGCTG GCGGATCAAG 1740  
 AAGACACTAC AAAAATTAG CAACAGTCCA GAGAAGCCTA AAGTGATTCA ATAGCCGAGG 1800  
 AGCACTGAT TCCTTCTGTC CTGAGGGGGT CTGGGGGGCT GGGGGGCGT GATGTGTGCC 1860  
 CTATCTGGGT AGAGGTAGTG TGAGTGTGGT GTGTGCTGGG GATGGGCAGC TGGCCTTGCC 1920  
 TGCTCGGCCC CAGGCCACCC CAGCCAAAAA TACAGCTGGG CTGAAACCTG ATCCCTTGCT 1980  
 GTCTGGCCTG CTCAAAGCGG CAGGCTCCTT GACGCTGGC TCTCTCCCA CCCCATGGC 2040  
 CAGCATGGTG CACCCCTTAC CACTCCCGGG ACAGGATGCA AAAGAGGCTC CAGAGTCAGA 2100  
 GTGCCAAGCC AGGGAATCCC AGTCCAGAC TCAGAGCCCG GGCTTGCACT TGGCCCCCTG 2160  
 CCTTGTATCA ACCCCACTGC CCCACCAGAG CTGCCAGGCT GGCACAGGGC CCTGTCCAGC 2220  
 CCTTGGCACA CACTTCCCTG CCAGGCCCTCA GCCTTAGCA TAAGCTCCAG AGAGCCAGGG 2280  
 CCCATCAGTT TCTCTCTGTG GATTGTATC TCAGCTCCAT GATGCCCTTG GCTTCTGTG 2340  
 TCTCAACAA GAGTGACAGT TGCTGAATGT CAGCTGCTGG AGAGAGCTGG GGCCTGACTT 2400  
 ACTAGGGCAT TAAATCCTAA GAGGTCCTAC TGAGGTGTGG CAGGATCACA GGCCAGTGG 2460  
 AAAAGGGCAG GTGAGTGGG CAGGCCCCAG GACTTTCAGA TTAAGTGA GAATATCAG 2520  
 GGCAGCATG GCAGGGGAA GGTCAAGTGG TGTCAAGAGA CCCAGGCTG ACCCCGAGT 2580  
 TTTGCTCCAT GTGACAAAG CAGGCTGTG TCAGGACCTT TTTTCTTCT TTTTCTTCT 2640  
 TTTTCTTCTT GACACGGAGT TTCGCTCTTG TTGTCCAGGC TAGAGTGCAA TGGCATGATC 2700  
 CCAGCTCACC GCAACGTCTA CCTCCAGGT TCAAATCATT CTCTTGCTC AGACTCCOGA 2760  
 GTAGCTGGGA TTACAGGCAC ATGCCACCAT GCCTGGCTAA TTTGTATAT TTAGTAGAAA 2820  
 CAGGGTTTCA CATGCTGGT CATGCTGGT CTGCAACTCC TGACCTCAGG TGTTCACCT 2880  
 ACCTCAGCCT CCCAAGTGC TGGGTTTACA GGTGTGAGCC ATCGCGCCTG GCCAGGACCT 2940  
 TTGTTCTTA TCTACATATT GGAAGATTG GTCTGATGT CCTTGGAGG TTTCTTAGCT 3000  
 CTAGTTCTCT CATGCTTCA GCTATATCAC AGCTAACTTC YTCAGTCTCA TCTATTCTT 3060  
 ATGCTCCAGC CCTGGCAAT TTGCTCAAG ATGGGGGTTT GAAATAAAT TTACTTGACT 3120  
 CAAGGAGTGT CTGGAGCAC TCTAGTCTA AGTCTGCAAG CTCCAGTTCT TGCCTAAAAC 3180  
 CATGCCAGTG GCCACCTTGT GGCTCAGACA GCTCTGGGCC TTTTGACCAC AAGCCAGCCC 3240  
 CTGCGCCTCT CTGTGGCATA GTCTTCTCTG CCCAGGACT GCAGGGGGCG TTCTTCAAG 3300  
 GCTTCAAGG CTCAAAAGAA ATTTGGCTCC ATCCAAGAG GCTCCAGCTC CCTACTGGC 3360  
 CCTGGCTTC AGGCCACAC CCTGGGCCA GGCSCAGAGA GTGTGTCTCA GGAGAAATTA 3420  
 ATGGGCTCTA GAGAGACACA CAGAAAGTTT GGGCATTTG GAAATTTTCA AGGRTGTATG 3480

5  
 10  
 15

TATGGYTCAC GTATGGWACA GGTGTGCTCTG GTCCYKGGGT GCAGGGAAGT GGGCTGCAGG 3540  
 GAAGTGGATT GGAGGGGAGC TTGAGGAATA TAAGGAGCGG GGTGAGAGAC TCAGGCTATG 3600  
 GACAAGGACA GCCCAAGGT TGGGAAGACC TGGCCTTAGT CGTCCCTCAGC CTAGGGCAGG 3660  
 CAGTGAAGA AAGCTCTCCC CGCTCCTGCT GTAATGACCC AGAGTAGCCT CCCAGGCCG 3720  
 GCATCTTATG TGTGTCTTCC ACCATCTCTCA TGGTGGCACT TTTCTAGGCC TGTCTCCAG 3780  
 CATTGTGCAA GGCTCGGAAG AGAACCAGGA AGTGAAACTG GGTGAAACA GAAAGCTCAA 3840  
 TGGATGGGCT AGGTTCACAG ATCATTAGGG CAGAGTTTGC ACGTCTCTG GTTCACTGGG 3900  
 AATCCACCCA GCCCAGCAAT CATCTCCCTC TTTGAAGGAT TTTWATTCT ACTGGGTTTT 3960  
 GGAACAAACT CTTGCTGAGA CCCCACAGCC AGAAACTGAA AGCAGCAGCT CCCCAAAGCC 4020  
 TGGAAATCC CTAAGAGAAG GCCTGGGGGA MAGGAATGG AGTGACAGGG GACAGGTAGA 4080  
 GAGAAGGGG CCCAATGGCC AGGAGTGAA GGAGGTGGCG TTGCTGAGAG CAGTCTGCAC 4140  
 ATGCTTCTGT CTGAGTGCAG GAAGGTGTTT CAGGGTGGAA ATTACACTTC TCGTACCTGG 4200  
 AGACGCTGTT TGTGGGAGCA CTGGGCTCAT GCCTGGCACA CAATAGGTCT GCAATAAACC 4260  
 ATGGTTAAAT CCTGAAAAA AAAAAAAA

Seq ID NO: 333 Protein sequence  
 Protein Accession #: NP\_000011

20  
 25  
 30

1 11 21 31 41 51  
 MTLGSPRKGL LMLLMALVTQ GDPVKPSRGP LVTCTCESPH CKGPTCRGAW CTVVLVREEG 60  
 RHPQEHRCGG NLHRELRCRGR PTEFVNHVCC DSHLCNHVVS LVLEATQPPS EQPGTDGOLA 120  
 LILGPVLALL ALVALGVGLL WVVRRRQEKQ RGLHSELGES SLILKASEQG DTMGLDLLDS 180  
 DCTTSGSGSL PFLVQRTVAR QVALVECVGK GRYGEVVRGL WHGESVAVKI PSSRDEQSWF 240  
 RETEYNTVL LREDNILGFI ASDMTSRNSS TQLWLITHYH EHGSLYDFLO RQTLPEPLAL 300  
 RLAVSAACGL AHLHVEIFGT QGKPAIAHRD FKSRLVVKSS NLQCCIALDLG LAVMHSQGS 360  
 YLDIGNPRV GTRKYMPEV LDEQIRTDCE ESYKWDINA FGLVLWEIAR RTIVNGIVED 420  
 YRPPFYDVVP NDPSFEDMKK VVCVDQQTPT IPNRLAADPV LSGLAQMRE CHYPNPSARL 480  
 TALRIKKTLO KISNSPERPK VIQ

Seq ID NO: 334 DNA sequence  
 Nucleic Acid Accession #: NM\_004126.1  
 Coding sequence: 108-329

35  
 40  
 45

1 11 21 31 41 51  
 GGCACGAGCT CGTCCCGGCC TTCAGTTGTT TCGGGACGCG CCGAGCTTGG CCGCTCTTCC 60  
 AGCGGCTCCG CTGCCAGAGC TAGCCCGAGC CCGGTTCTGG GCGGAAATG CCGTCCCTTC 120  
 ACATCGAAGA TTTGCCAGAG AAGGAAAAAC TGAAATGGA AGTTGAGCAG CTTGCGAAAG 180  
 AAGTGAAGTT GCAGAGACAA CAAGTGCTTA AATGTTCTGA AGAAATAAAG AACTATATTG 240  
 AAGAACGTTT TGGAGAGGAT CCTCTAGTAA AGGGAATTCC AGAAGACAAG AACCCCTTTA 300  
 AAGAAAAAGG CAGCTGTGTT ATTTCATAAA TAACITGGGA GAAACTGCAT CCTAAGTGGA 360  
 AGAACTAGTT TGTTTTAGTT TTCCAGATA AAACCAACAT GCTTTTAAAG GAAGGAAGAA 420  
 TGAAATTAAG AGGAGACTTT CTTAAGCACC ATATAGATAG GGTATGTAT AAAAGCATAT 480  
 GTGCTACTCA TCTTTGCTCA CTATGCAGTC TTTTAAAGA GAGCAGAGAG TATCAGATGT 540  
 ACAATTATGG AATTAAGAAC ATTACTTGAG CATGACACTT CTTTCAGTAT ATTGCTTGAT 600  
 GCTTCAAATA AAGTTTTGTC TT

Seq ID NO: 335 Protein sequence  
 Protein Accession #: NP\_004117.1

55

1 11 21 31 41 51  
 MPALHIEDLP EKEKLEMEVE QLRKEVKLQR QVSKCSEBI KNYIEERSGE DPLVKGIPED 60  
 KNPFKEKGSC VIS

Seq ID NO: 336 DNA sequence  
 Nucleic Acid Accession #: NM\_005795  
 Coding sequence: 555-1940

65  
 70  
 75  
 80  
 85

1 11 21 31 41 51  
 GCACGAGGGA ACAACCTCTC TCTCTSCAGC AGAGAGTGTC ACCTCCTGCT TTAGGACCAT 60  
 CAAGCTCTGC TAACCTGAATC TCATCCTAAT TGCAGGATCA CATTGCAAAG CTTTCACTCT 120  
 TTCCACCTTT GCTTGTGGGT AAATCTCTTC TGGGGAATCT CAGAAAGTAA AGTTCCATCC 180  
 TGAGAATATT TCACAAAGAA TTTCCCTAAG AGCTGGACTG GGTCTTGACC CCGGAATTT 240  
 AAGAAATCTT TAAAGACAAT GTCAAATATG ATCCAAGAGA AAATGTGATT TGAGTCTGGA 300  
 GACAATTGTG CATATGCTCT AATAATAAAA ACCCATACTA GCCTATAGAA AACAAATATT 360  
 GAATAATAAA AACCCATACT AGCCTATAGA AAACAATATT TGAAGATTG CTACCACTAA 420  
 AAAGAAACT ACTACAACCT GACAAGACTG CTGCAAACTT CAATTGGTCA CCACAACCTG 480  
 ACAAGTTGTC TATAAAACAA GATTGCTACA ACTTCTAGTT TATGTTATAC AGCATATTT 540  
 ATTTGGGCTT AATGATGGAG AAAAAGTGTA CCTGTATT TTCTGTTCTC TTGCTTTTTT 600  
 TTATGATTCT TGTACAGCA GAATTAGAAG AGAGTCTTGA GGACTCAATT CAGTTGGGAG 660  
 TTACTAGAAA TAAATCATG ACAGCTCAAT ATGAATGTTA CCAAAAGATT ATGCAAGACC 720  
 CCATTCAACA AGCAGAAGGC GTTACTGCA ACAGAACCCT GGATGGATGG CTCTGCTGGA 780  
 ACGATGTTGC AGCAGGAACCT GAATCAATGC AGCTCTGCCG TGATTACTTT CAGGACTTTG 840  
 ATCCATCAGA AAAAGTTACA AAGATCTGTG ACCAAGATGG AAATGTTT AGACATCCAG 900  
 CAAGCAACAG AACATGGACA AATTATACC AGTGTAAATG TAACACCCAC GAGAAAGTGA 960  
 AGACTGCACT AAATTTGTTT TACCTGACCA TAATTGGACA CGGATTGTCT ATTGCATCAC 1020  
 TGCTTATCTC GCTTGGCATA TTCTTTTATT TCAAGAGCCT AAGTTGCCAA AGGATTACCT 1080  
 TACACAAAAA TCTGTTCTTC TCATTGTTT GTAACTCTGT TGTAACTATC ATTCACCTCA 1140  
 CTGCAAGTGC CAACAACCTG GCCTTAGTAG CCACAAATCC TGTAGTTGCA AAAGTGTCCC 1200  
 AGTTTACTTA TCTTACCTG ATGGGCTGTA ATTACTTTG GATGCTCTGT GAAGGCATT 1260  
 ACCTACACAC ACTCATTTGT GTGGCCGTGT TTGCAGAGAA GCAACATTTA ATGTGGTATT 1320  
 ATTTTCTTGG CTGGGGATT TCCACTGATT CTGCTGTAT ACATGCCATT GCTAGAAGCT 1380



TATATTACAA TGACAAITTC TGGATCAGTT CTGATACCCA TCTCCTCTAC ATTATCCATG 1440  
 GCCCAATTTC TGCTGCTTTA CTGGTGAATC TTTTTCCTT GTTAAATATT GTACGCGTTC 1500  
 TCATCACCAG GTTAAAGATT ACACACCAAG CGGAATCCAA TCTGTACATG AAAGCTGTGA 1560  
 GAGCTACTCT TATCTTGGTG CCATTGCTTG GCATTGAATT TGTGCTGATT CCATGGCCGAC 1620  
 CTGAAGGAAA GATTGCAGAG GAGGTATATG ACTACATCAT GCACATCCTT ATGCACTTCC 1680  
 AGGGTCTTTT GGTCTCTACC ATTTCTCTGT TCTTTAATGG AGAGGTTCAG GCAATTCCTG 1740  
 GAAGAACTCG GAATCAATAC AAAATCCAAAT TTGGAAACAG CTTTCCCAAC TCAGAAGCTC 1800  
 TTCGTAGTGC GTCTTACACA GTGTCAACAA TCAGTGTAGG TCCAGGTTAT AGTCATGACT 1860  
 GTCCTAGTGA ACACCTTAAAT GGAAGAAAGCA TCCATGATAT TGAAATGTT CTCTTAAAC 1920  
 CAGAAATTT ATATAATTGA AATAGAAGG ATGGTGTCT CACTGTTTGG TGCTTCTCCT 1980  
 AACTCAAGGA CTTGACCCCA TGACTCTGTA GCCAGAAGAC TTCAATATTA AATGACTTTG 2040  
 GGGAAATGTA TAAAGAAGAG CCTTCACATG AAATTAGTAG TGTGTTGATA AGAGTGTAAAC 2100  
 ATCCAGCTCT ATGTGGGAAA AAAGAAATCC TGGTTGTAA TGTGTTGAG TAAATACTCC 2160  
 CACTATGCTT ACATGTAGCG TACTAACCTG ACATCACCAA GTGTGGAATT GGAGAAAGC 2220  
 ACAATCAACT TTTCTGAGCT GGTGAAGCC AGTCCAGCA CACCATTGAT GAATTCAAAC 2280  
 AAATGGCTGT AAAACTTAAAC ATACATGTTG GGCATGATTC TACCCTTATT CCCCCAAGA 2340  
 GACCTAGCTA AGGTCTATAA ACATGAAGGG AAAATTAGCT TTAGTTTITA AAATCTTTTA 2400  
 TCCCATCTTG ATTTGGGGCAG TTGACTTTTT TTTTTCCTCA GAGTGCCGTA GTCCCTTTTG 2460  
 TAATACCTCT CTCAATGGA CAATACCAGA AGTGAATTAT CCGTCTGGC TTTCTTTCT 2520  
 CTATGAAAG CAATGAGTA CAATTGTAT GATCTACTCA TTTGCTGACA CATCAGTTAT 2580  
 ATCTGTGGC ATATCCATG TGGAACTGG ATGAACAGGA TGTATAATAT GCAATCTTAC 2640  
 TTCTATATCA TTAGGAAAAC ATCTAGTTG ATGCTACAAA ACACCTTGTC AACCTCTTCC 2700  
 TGTCTTACCA AACAGTGGGA GGGAAATCCT AGCTGTAAT ATAAATTTTG CCCTTCCATT 2760  
 TCTACTGTAT AAACAAATTA GCAATCATT TATATAAGA AAATCAATGA AGGATTTCTT 2820  
 ATTTTCTTGG AATTTGTAA AAAGAAATG TGAAATATGA GCTGTAAAT ACTCCATTAT 2880  
 TTTATTTTAT AGTCTCAAT CAAATACATA CAACTATGT AATTTTAA GCAATATAT 2940  
 AATGCAACAA TGTGTATG TTAATATCTG ATACTGTATC TGGGCTGATT TTTTAAATAA 3000  
 AATAGAGTCT GGAATGCT

Seq ID NO: 337 protein sequence  
 Protein Accession #: NP\_005786.1

1 11 21 31 41 51  
 MEKKCTLYFL VLLPFFMILV TABLESPED SIQLGVTRNK IMTAQYECYQ KIMQDPIQQA 60  
 EGVYQNRWMD GMLCWNVDAA GTESMQLCPD YFQDFDPSEK VTKICDQDGN WFRHPASNRT 120  
 WTNVYQCNVN THEKVKITALN LPYLTIIIGEG LSIALLLISL GIFFYFKSLR CQRITLHKNL 180  
 FFSFVCNSVV TIHLTAVAN NQALVATNPV SCKVSQPIHL YLNGCNYFMW LCEGIVLHLD 240  
 IVVAVFAEKQ HLMWYFLLGW GFPLIPACIH AIARSLYYND NCHISSDTHL LYIIHGPICA 300  
 ALLVNLFFLL NIVRVLITKL KVTHQAESNL YMKAVRATLI LVPLLGIEFV LIPWRPEGKI 360  
 AEEVDYIMH ILMHQGLLV STIFCPFNGE VQAILRRNWN QYKIQPCNSF SNSEALRSAS 420  
 YTVSTISDGP GYSHDCPSEH LMGKSIHDIH NVLLKPEPLY N

Seq ID NO: 338 DNA sequence  
 Nucleic Acid Accession #: NM\_001795  
 Coding sequence: 25-2379

1 11 21 31 41 51  
 GCACGATCTG TTCCTCTGG GAAGATGCAG AGGCTCATGA TGCTCCTCGC CACATCGGGC 60  
 GCGTGGCTGG GCGTCTGGC AGTGGCAGCA GTGGCAGCAG CAGGTGCTAA CCTGCCCCAA 120  
 CGGGACACCC ACAGCCTGCT GCCCACCCAC CGGCGCCAAA AGAGAGATTG GATTGTGAAC 180  
 CAGATGCACA TTGATGAAGA GAAAAACACC TCACTTCCCC ATCATGTAGG CAAGATCAAG 240  
 TCAAGCGTGA GTGCAAGAA TGCCAAAGTAC CTGCTCAAAG GAGAATATGT GGGCAAGGTC 300  
 TTCGGGTGCG ATGCAGAGAG AGGAGACGCTG TTCGCCATTG AGAGGCTGGA CCGGGAGAAT 360  
 ATCTCAGAGT ACCACCTCAC TGCTGTCAAT GTGACAAGG ACACCTGGTA AAACCTGGAG 420  
 ACTCCTTCCA GCTTCACCAT CAAAGTTCAT GACGTGAAGC AACTAGGCC TGTGTTCAAG 480  
 CATCGGTTGT TGAATGCGTC CGTGCTGAG TCGTGGCTG TGGGACCTC AGTCATCTCT 540  
 GTGACAGCAG TGAATGCAGA CGACCCCACT GTGGGAGACC ACGCCTCTGT CATGTACCAA 600  
 ATCCTGAAGG GGAAAGAGTA TTTTGCCATC GATAATTCTG GACGTATTAT CACAATAACG 660  
 AAAAGCTTGG ACCGAGAGAA GCAGGCCAGG TATGAGATCG TGGTGAAGC GCGAGATGCC 720  
 CAGGGCCTCC GGGGGGACTC GGGCAAGGCC ACCGTGCTGG TCACTCTGCA AGACATCAAT 780  
 GACAACTTCC CTTCTTTCAC CCAGACCAAG TACACATTTG TCGTGGCTGA AGACACCGCT 840  
 GTGGGCACTT CTGTGGGCTC TCTGTTTGT GAGGACCCAG ATGAGCCCCA GAACCGGATG 900  
 ACCAAGTACA GCATCTTGG GGGCGACTAC CAGGACGCTT TCACCATGGA GACAAACCCC 960  
 GCCCACAAAG AGGGCATCAT CAAGCCCATG AAGCCTCTGG ATTATGAATA CATCCAGCAA 1020  
 TACAGCTTCA TGTGAGAGG CACAGACCCC ACCATCGACC TCCGATACAT GAGCCTCCCC 1080  
 GCGGGAAACA GAGCCAGGT CATTATCAAC ATCAGAGATG TGGACGAGCC CCCCATTTC 1140  
 CAGCAGCCTT TCTACCACTT CCAGCTGAAG GAAAACCAGA AGAAGCCTCT GATTGGCACA 1200  
 GTGCTGGCCA TGGACCTTGA TGGGCTAGG CATAGCATTG GATACTCCAT CCGCAGGACC 1260  
 AGTGACAGAG GCCAGTTCTT CCGAGTCACA AAAAAGGGGG ACATTACAA TGAGAAAGAA 1320  
 CTGGACAGAG AAGTCTACCC CTGGTATAAC CTGACTGTGG AGGCCAAAGA ACTGGATTCC 1380  
 ACTGGAACCC CCACAGGAAA AGAATCCATT GTGCAAGTCC ACATTGAAGT TTTGGATGAG 1440  
 AATGACAAAT CCCCAGGATT TGCCAGCCCC TACCAGCCCC AAGTGTGTGA GAACGCTGTC 1500  
 CATGCCACGC TGGTCTGCA GATCTCGCA ATAGACAAGG ACATAACRCC ACGAAACGTG 1560  
 AAGTTCAAAT TCACCTTGAA TACTGAGAAC AACTTTACCC TCACGGATAA TCACGATAAC 1620  
 ACGGCCAACA TCACAGTCAA GTATGGGCAG TTTGACCGGG AGCATAACAA GGTCCAATTC 1680  
 CTACCCGTGG TCATCTCAGA CAATGGGATG CCAAGTCGCA CGGGCACAG CACGCTGACC 1740  
 GTGGCGTGT GCAAGTCAA CGAGCAGGGC GAGTTCACTT TCTGCGAGGA TATGGCCGCC 1800  
 CAGGTGGGCG TGAGCATCCA GGCAGTGGTA GCCATCTTAC TCTGCATCCT CACCATCACA 1860  
 GTGATCACCC TGCTCATCTT CCGTGGGCGG CGGCTCCGGA AGCAGGCCCG CGGCAACGGC 1920  
 AAGAGCGTGC CGAGATCCA CGAGCAGCTG CTCACTACG ACGAGGAGGG CGGCGCGGAG 1980  
 ATGGACACCA CCAGCTACGA TGTGTCGGTG CTCACTCGG TGGCGCGGG CGGGGCCAAG 2040  
 CCCCAGCGGG CCGGCTTGA CGCCCGGCTT TCCCTCTATG CGCAGGTGCA GAAGCCACOG 2100  
 AGGCAAGCGC CTGGGGCACA CGGAGGGCCC GGGAGATGG CAGCATGAT CGAGGTGAAG 2160  
 AAGGACGAGG CGGACACGA CGGCGACGCG CCCCCTACG ACACGCTGCA CATCTACGGC 2220  
 TACGAGGGCT CCGAGTCCAT AGCCGAGTCC CTGAGTCCCC TGGGCAACGA CTCATCCGAC 2280



TCTGACGTGG ATTACGACTT CCTTAAAGAC TGGGGACCCA GGTTTAAGAT GCTGGCTGAG 2340  
 CTGTACGGCT CGGACCCCGG GGAGGAGCTG CTGTATTAGG CGGCCGAGGT CACTCTGGGC 2400  
 CTGGGGGACC AAACCCCTCG CAGCCGAGGC CAGTCAGACT CCAGGCACCA CAGCCTCCAA 2460  
 AAATGGCAGT GACTCCCCAG CCCAGCAACC CTTCCTCGTG GGTCCAGAG ACCTCATCAG 2520  
 CCTTGGGATA GCAAACTCCA GGTTCCTGAA ATATCCAGGA ATATATGTCA GTGATGACTA 2580  
 TTCTCAAAAT CTGGCAAAAT CAGGCTGGTG TTCTGTCTGG GCTCAGACAT CCACATAACC 2640  
 CTGTCAACCA CAGACCGCGG TCTAACTCAA AGACTTCTCT TGGCTCCCCA AGGCTGCAAA 2700  
 GCAAAACAGA CTGTGTTTAA CTGCTGACAG GTCTTTTCTT AGGGTCCCTG AACGCCCTGG 2760  
 TAAGGCTGGT GAGGCTCTGG TGCCCTATCTG CCTGGAGGCA AAGGCTTGGG CAGCTTGACT 2820  
 TGTGGGGCAG GATTCTCTGC AGCCCATTCG CAAGGGAGAG TGACCATCAT GCCCTCTCTC 2880  
 GGGAGCCCTA GCCCTGCTCC AACTCCATAC TCCACTCCAA GTGCCCCACC ACTCCCCAAC 2940  
 CCTCTCTCAG GCCTGTCAAG AGGGAGGAAG GGGCCCCATG GCAGCTCTCT ACCTTGGGTC 3000  
 CTGAAGTGAC CTCACGTGCC TGCCATGCCA GTAAGTGTGC TGTACTGAGC ACTGAACCAC 3060  
 ATTCAGGGAA CTGCTTATTA AACCTTGAAG CAACTGTGAA TTCATTCTGG AGGGGCGAGT 3120  
 GAGATCAGGA GTGACAGATC ACAGGGTGAG GGCCACCTCC ACACCCACCC CCTCTGGAGA 3180  
 AGGCTTGGA GAGCTGAGAC CTGTCTTGA GACTCCTCAG CACCCCTCCA GTTTTGGCTG 3240  
 AGAAGGGGCA GATGTTCCCG GAGATCAGAA GAGCTCTCCC CTCTCTGCTC TCACCTGGTC 3300  
 GCCAATCCAT GCTCTCTTTC TTTTCTCTGT CTACTCTCTA TCCCTTGGTT TAGAGGAACC 3360  
 CAAGATGGG CCTTTAGCAA AACTGACAA GTCCAAACCC ACTCATGACT GCATGACGGA 3420  
 GCCGAGCATG TGTCTTTACA CTTGCTGTT GTCCATCTC AGGGAACCTG COCTCAGGCA 3480  
 CACCTTGAG CAGGAAGGCC CTGCCCTGCC CAACTCTGT GGTCAACCAT GCATCAITTC 3540  
 ACTGGAACGT TTCCTGCAA ACACACCTTG GAGAAGTGGC ATCAGTCAAC AGAGAGGGGC 3600  
 AGGGAAGGAG ACACCAAGCT CACCTTCTGT CATGGACCGA GGTTCCTCAT CTGGCAAAGC 3660  
 CCTCACACT GCAAGGATG GTAGATAACA CTGACTTGT TGTTTTAAAC AATAACTAGC 3720  
 TTCTTATAAT GATTTTTTTA CTAATGATAC TTACAAGTT CTAGCTCTCA CAGACATATA 3780  
 GAATAAGGGT TTTTGCATA TAAGCAGGTT GTTATTAGG TTAACAATAT TAATTGAGT 3840  
 TTTTATGTTG GAAAAACAA TCCTGTAAAC TTCTATTTC TATAATTGTA GTAATTGCTC 3900  
 TCACGATAAT GTCTATATAT TGGCCAAACT GGTGCATGAC AAGTACTGTA TTTTATTATA 3960  
 CCTAATAAAA GAAAAATCTT TAGCCTGGGC AACAAAAAAA

Seq ID NO: 339 Protein sequence  
 Protein Accession #: NP\_001786

1 11 21 31 41 51  
 MQRIMLLLAT SGACLGILLAV AAVAAAGANP AQRDTHSLLP TERRQKRDWI NNQMHIIEEK 60  
 NTSLSPHVVGK IKSSVSRRKNA KYLLKGEYVG KVFVRDAETG DVFAIERLDR ENISEYHLTA 120  
 VIVDKDTGEN LBTPSPFTIK VHDVNDNMFV FTRLEFNASV PRESSAVGTSV ISVTAVDADD 180  
 PTVGDHASFVM YQILKGKEYF AIDNSGRIIT ITKSLDREKQ ARYEIVVEAR DAQGLRGDSG 240  
 DYTALVLTLDQ INDNFPFPTQ TKTPTVPVED TRVGTSGVSL FVEDPDEPQN RMTKYSILRG 300  
 DYQDAFTLET NPAHNEGIIK PMKPLDYEYI QQYSFIVEAT DPTIDLRYMS PPAGNRAQVI 360  
 INITDVEDEP IQPQPFYHFQ LKENGKFLI GTVLAMPDPA ARHSIGYSIR RTSKDGQFFR 420  
 VTRKGDVYNE KELDREVVPW YNLTVAEKEL DSTGTPTGKE SIVQVHIEVL DENDNAPEFA 480  
 KPYQPKVCEN AVHQQLVLQI SAIDKDIPTP NVKFKPTLNT ENNFTLTDNH DNTANITVKY 540  
 GQFDREHTRV HPLFPVVISDN GMPSTRGTST LTVAVCKCNE QGEPTFCEDM AAQVGVSIAQ 600  
 VVAILLCLLT ITVITLLIPL RRRLRKQARA HGKSVPEIHE QLVTYDEEGG GEMDITTSYDV 660  
 SVLNSVRRGG AKPFRPALDA RPSLYAQVQK PPRHAPGAHG GPGEMAAMIE VKKDEADHDG 720  
 DGPPYDTLHI YGYEGSESLA ESSLGLTDS SDSVDVDFL NDWGRPRFML AELYGSDPRE 780  
 ELLY

Seq ID NO: 340 DNA sequence  
 Nucleic Acid Accession #: NM\_003088  
 Coding sequence: 112-1593

1 11 21 31 41 51  
 GCGGAGGGTG CGTGCGGGCC GCGGCAGCCG AACAAAGGAG CAGGGGCGCC GCGCAGGGA 60  
 CCGGCCACCC ACTCCCGGGG GCGCGGCAGC GGCTCTCGT CTACTGCCAC CATGACCGCC 120  
 AACGGCACAG CCGAGGCGGT GCAGATCCAG TTCGGCTCTA TCAACTGCGG CAACTAGTAC 180  
 CTGACGGCCG AGGCGTTCGG GTTCAAGSTG AACCGTCCG CCAGCAGCCT GAAGAAGAAG 240  
 CAGATCTGGA CGCTGGAGCA GCCCCTGAC GAGGCGGGCA GCGCGGCCGT GTGCTCGCG 300  
 AGCCACCTGG GCCGTACTCT GCGCGCGGAC AAGGAACGCA ACGTGACCTG CGAGCGCGAG 360  
 GTGCCCGGTC CCGACTGCCG TTTCTCATC GTGGCGCAGC ACGACGGTCG CTGGTCGCTG 420  
 CAGTCCGAGG CGCACCGCGC CTACTTCGCG GGCACCGAGG ACCGCTGTCT CTGCTTCGCG 480  
 CAGACGGTGT CCCCAGCGGA GAAGTGGAGC GTGCACATCG CCATGCACCC TCAGGTCAAC 540  
 ATCTACAGTG TCACCGGTAA GCGCTAGCGG CACTGAGGCG CCGCGCCGCG CGACGAGATC 600  
 GCCGTGGACC GCGACGTGCC CTGGGGCGTC GACTCGCTCA TCACCTCTCG CTTCAGGAC 660  
 CAGCGCTACA GCGTGACAGC CGCCGACCA CCGCTCTCTG GCCACGACGG GCGCCTGGTG 720  
 GCGCGCCCGG AGCGGCGCAC TGGCTACACG CTGGAGTTCC GCTCCGGCAA GGTGGCCTTC 780  
 AAGGCACTGG AGGGCGGTTA CCTGGGCGCG TCGGGGCCCA GCGGCACGCT CAGGCGGGC 840  
 AAGGCCACCA AGGTGGGCAA GGACGAGCTC TTGTCTCTGG AGCAGAGCTG CGCCAGGTC 900  
 GTGCTGCACG CGGCCAACGA GAGGAACGTG TCCACGCGCC AGGGTATGGA CTGTCTGCC 960  
 AATCAGGACG AGGAGACCGA CCAGGAGACC TTCCAGCTGG AGATGACCG CGACACCAAA 1020  
 AAGTGTGCTT TCGTACCCA CACGGCAAG TACTGGAAGC TGAAGGCCAC CGGGGCGGTG 1080  
 CAGTCCACCG CCTCCAGCAA GAATGCCAGC TGCTACTTTG ACATGAGTG GGTGACCGG 1140  
 CGCATCACAC TGAGGGCGTC CAATGGCAAG TTTGTGACCT CCAAGAAGAA TGGCGAGCTG 1200  
 GCGGCTCGG TGAGACAGC AGGGGACTCA GAGCTCTTCC TCATGAAGCT CATCAACCGC 1260  
 CCCATCATCG TGTTCGCGG GGAGCATGGC TTCACTCGCT GCGCAAGGT CACGGGCACC 1320  
 CTGGAAGCCA ACGCTCCAG CTATGACGTC TTCCAGCTGG AGTTCAACGA TGGCGCTTAC 1380  
 AACATCAAAG ACTCCACAGG CAAATACTGG ACGTGGGCA GTGACTCCGC GTTCAACAGC 1440  
 AGCGGCGACA CTCTGTGGA CTCTTCTTCT GAGTCTCTGC ACTATAACAA GGTGGCCATC 1500  
 AAGGTGGGCG GCGCTACTCT GAAGGGCGAC CACGAGGCG TCCTGAAGGC CTCGGCGGAA 1560  
 ACCGTGGAAC CCGCTCGCT CTGGGAGTAC TAGGGCGCGC CCGTCTTCC CCGCCCTTGC 1620  
 CCACATGGCG GCTCTGCGCA ACCCTCCCTG CTAACCCCTT CTCGGCCAGG TGGGCTCCAG 1680  
 GCGGAGAGCG AAGCCCCCTT GCCTTCAAA CTGGAACCC CAGAGAAAC GGTGCCCCCA 1740  
 CCTGTGCGCC CTATGGAATC CCACTCTTCC CTTCCGCGCG GGTCCCTTAC TCCCTCGGG 1800

TCAGCGGCTG CGGCCCTGGCC CTGGGAGGGA TTTCAGATGC CCTTGCCCTC TTGTCTGCCA 1860  
 CGGGGGAGT CTGGACCTTC TTCTTCTGA CCTCAGACGG CTCTGAGCCT TATTCTCTCTG 1920  
 GAAGCGGCTA AGGAGCGGTT GGGGGCTGGG AGCCCTGGGC GTGTAGTGTA ACTGGAATCT 1980  
 TTTCCTCTC CCAGCCACCT CCTCCAGGCC CCCAGGAGA GCTGGGCACA TGTCCCAAGC 2040  
 CTGTCACTGG CCTCCCTGG TGCACTGTCC CGGAAACCCC TGCTTGGGAA GGGAAAGCTGT 2100  
 CGGGAGGGCT AGGACTGACC CTTGTGTGT TTTTGTGGT GGTGGCTGGA AACAGCCCT 2160  
 CTCCACGCTG GGAGAGGCTC AGCCTGGCTC CCTTCCCTGG AGCGGCAGGG CGTGACGGCC 2220  
 ACAGGGCTCG CCGCTGTCAC GTTCTGCCAA GGTGTGTGTG GCGGGCGGGT AGGGGTGTGG 2280  
 GGGCGCTCTT CCTCTGTCT CTTCTCTTC ACCCTAGCCT GACTGGAAGC AGAAAATGAC 2340  
 CAAATCAGTA TTTTCTTAA TGAATATTA TTGCTGGAGG CGTCCAGGC AAGCCTGGCT 2400  
 GTAGTAGCGA GTGATCTGGC GGGGGGCTC TCAGCACCTT CCCAGGGGG TGCATCTCAG 2460  
 CCCCCTCTTT CGTCTCTCC CGTCCAGCC CAGCCCTGGG CTTGGGCTGC CGACACCTGG 2520  
 GCGAGAGCCC CTGCTGTGAT TGTGTCTCC TGGGCTCTCC GGGTGGATGA AGCCAGGCGT 2580  
 CGCCCCCTCC GCGAGCCCTG GGGTGAGCG CCGGGGCCCC CCGTCTGCCA GCCTCCCCCG 2640  
 TCCCCAATC GCATCTCACT CTGGGTGTCT TGGTCTTTA TTTTGTGTA GTGTCAATTG 2700  
 TATAACTCTA AACGCCCATG ATAGTAGCTT CAACTGGAA ATAGCGAAAT AAAATAACTC 2760  
 AGTCTGC

Seq ID NO: 341 Protein sequence  
 Protein Accession #: NP\_003079

1 11 21 31 41 51  
 MTANGTAERV QIQFLINCG NKYLTAEPF FKNVASASSL KKKQIWTLEQ PFDEAGSAAV 60  
 CLRSHLGRYL AADKDNVTC EREVGPDCR FLIVAHDDGR WSLQSEAHRR YFGTEDRLS 120  
 CFAQTVPAP KWSVHIAHP QVNIYSVTRK RYAHLSARPA DELAVDRDP WGVDSLITLA 180  
 FQDQRYSVQT ADERFLRHDG RLVARPEPAT GYTLEFRSGK VAFRDCBGRY LAPSGPSGTL 240  
 KAGKATKVGK DELFALEQSC AQVVLQAANE RNVSTRQGM DLSANQDEETD QETPQLEIDR 300  
 DTKKCAFRTH TGRYWTLAT GGVQSTASSK NASCYFDIEW RDRRITLRAS NGKFPVTSKN 360  
 GQLAABVETA GDELFLMKL INRPIIVFRG EHGFIGCRKV TGTLDANRSS YDVPQLEFND 420  
 GAYNIKDSGT KYVTGSDSA VTSSGDTFVD FPFECDDYNK VAIKVGGRYL RGDHAGVLKA 480  
 SAETVDPASL WEY

Seq ID NO: 342 DNA sequence  
 Nucleic Acid Accession #: FGENESH predicted  
 Coding sequence: 660..1705

1 11 21 31 41 51  
 CGCTCCGCAC ACATTTCCTG TCGCGGCCTA AGGGAAACTG TTGGCCGCTG GCGCCGCGGG 60  
 GGGATTCTTG GCACTTGGGG GGTCCGTCTG GAGCGAGGGC GGAGGGGAAG GGAGGGGGAA 120  
 CCGGGTTGGG GAAGCCAGCT GTAGAGGGCG GTGACCGCGC TCCAGACACA GCTCTGCGTC 180  
 CTGAGCGGGG ACAGATCCAA GTTGGAGACA GCTCTGCGTG CCGGGCCCTCA GAGAATGAGG 240  
 CCGGCGTTGG CCTGTGCTCT CTTCTGGCAG GCGCTCTGGC CCGGGCCGGG CGGCGGCGAA 300  
 CACCCCACTG CCGACCGTGG TGGCTGCTGG GCTCTGCGGG CTTGCTACAG CTTGCACCTC 360  
 GCTACCATGA AGCGGCGAGG GCGCGAGGAG GCCTGCATCC TCGGAGGTGG GCGGCTCAGC 420  
 ACCGTGCGTG CCGGGCGCGA GCTGCGCGCT GTGCTGCGCG TCGTGGGGG AGGCCAGGGG 480  
 CCGGAGGGGG GCTCCAAAGA CTTGCTGTTC TGGGTGCGAC TGGAGCGCAG CGGTTCCTCC 540  
 TGCACCTTGG AGAACGAGCC TTTCGCGGGT TTCTCTGGC TGTCTCTCGA CCGCGCGCGT 600  
 CTCGAAAGCG ACACGCTGCA GTGGGTGGAG GAGCCCCAAC GCTCCTGCAC CGCGCGGAGA 660  
 TGGCGGGTAT TCCAGGCGAC CGGTGGGGTC GAGCCCGCAG CTGGAAGGAG ATGCGATGCC 720  
 ACCCTGCGCG CAACGCTTAC CTGTGCAAGT ACCAGTTTGA GGTCTTGTGT CTTGCGCGCG 780  
 GCGCGGGGGG CCGCTCTAAC TTGAGCTATC GCGCGCCCTT CCAGCTGCAC AGCGCGGCTC 840  
 TGGACTTCAG TCCACCTGGG ACCGAGGTGA GTGCGCTCTG CCGGGGACAG CTTCCGATCT 900  
 CAGTTACTTG CATCGCGGAC GAAATCGGGG CTCGCTGGGA CAAACTCTCG GCGATGTGT 960  
 TGTGTCCCTG CCGCGGGAGG TACCTCCGTG CTGGCAAATG CGCAGAGCTC CCTAATCGCC 1020  
 TAGAGCACTT GCGAGGCTTT GCTTGGGAAT GTGCTACGGG CTTGAGCTG GGGAGGAGCG 1080  
 GCGCTCTTGT TGTGACCACT GGGGAAGGAC AGCCGACCTT TGGGGGGACC GGGGTGCCCA 1140  
 CCAGCGCGCG CCGGCGCACT GCAACGAGCC CCGTGCAGCA GAGAACATGG CCAATCAGGG 1200  
 TCGAGAGGAA GCTGGGAGAG ACACCACTTG TCCCTGAACA AGACAATTCA GTAACATCTA 1260  
 TTCTTGAGAT TCTCTGATGG GGATCACAGA GCAAGATGTC TACCCCTCAA ATGTCCCTTC 1320  
 AAGCCGAGTC AAAGGCCACT ATCAGCCCAT CAGGGAGCGT GATTTCCTAA TTTAATTCTA 1380  
 CGACTTCTCT TGCCACTCCT CAGGCTTTGG ACTCCTCTCT TCGCGTGGTC TTCATATTGT 1440  
 TGAGCACAGC AGTAGTAGTG TTGGTGATCT TGACCATGAC AGTACTGGGG CTTGTCAAGC 1500  
 TCTGCTTTCA CGAAAGCCCC TCTTCCAGC CAAGGAAGGA GTCTATGGGC CCGCGGGGCC 1560  
 TGGAGAGTGA TCTTGAGCCC GCTGCTTTGG GCTCCAGTTC TGACATTTCG CAAACCAATG 1620  
 GGGTGAAAGT CCGGAGCTGT GATCTGCGGG ACAGAGCAGA GGGTGCTCTG CTGGCGGAGT 1680  
 CCCCCTCTGG CTCTAGTGAT GCATAG

Seq ID NO: 343 Protein sequence  
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51  
 MGKDFMTKTP KAFATKAKID KWDLIKLSF CTAKETIIRV NSQPTDWQKT FAIYPSDKGV 60  
 IARIYKLEQ IYKKKPKTK LRTHFLSRPK GNCNPLGPRG DSWQLGGPSG ARABCKGGGT 120  
 GLGKPAVEGG DRAPDTALRP RAGQIQVSS SAOCASENEA GVREVPPLAG ALARAGRRRT 180  
 PHCRPCLLLG LGGLLQAPPR YHEAAGRGGR LHPARWGAQH RACGRRAARC ARAPAGRPR 240  
 RRGLQRPAVL GRTGAQAPPL HPGERAFAGP LLAVLRPRRS RKRHAHVGGG APTLLHRAEM 300  
 RGTGPHRWGR ARSWKEMRCH LRANGYLCKY QFEVLCPAPR PGAASNLVSR APFQLHSAAL 360  
 DFPSPGTEVS ALCRGQLPIS VTCIADEIGA RWDKLSGDVL CPCFGYLRRA GKCAELPNCL 420  
 DDLGGFACBC ATGFELGKDG RSCVTSGEQG PTLGGTGVPRT RRPPATATSP VPORTWPIRV 480  
 DEKLGETPLV PEQDNSTSI PEIPRWGSQS TMSLTQMSLQ AESKATITPS GSVISKPNST 540  
 TSSATPQAFD SSSAVVIFV STAVVVLVIL TMTVLGLVKL CFHESPSSQP RKBSMGPPGL 600  
 ESDPEPAALG SSSARCTNNG VKVGDCLDRD RAEGALLAES PLGSSDA

Seq ID NO: 344 DNA sequence  
Nucleic Acid Accession #: NM\_012072  
Coding sequence: 149-2107

5  
0  
5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

```

1.      11      21      31      41      51
|        |        |        |        |        |
AAAGCCCTCA GCCTTGTGT CCTTCTCTGC GCCGGAGTGG CTGCAGCTCA CCCCTCAGCT 60
CCCCTTGGGG CCCAGCTGGG AGCCGAGATA GAAGCTCCTG TCGCGCTCGG GCTTCTCGCC 120
TCCCGCAGAG GGCACACAG AGACCGGAT GCCACCTCC ATGGGCTGCG TGCTGCTGCT 180
GCTGCTGCTC CTGACCCAGC CCGGGGCGGG GACGGGAGCT GACACGGAGG CGGTGGTCTG 240
CGTGGGGACC GCCTGCTACA CGGCCACTC GGGCAAGCTG AGCGCTGCG AGGCCAGAA 300
CCACTGCAAC CAGAACGGGG GCAACCTGCG CACTGTGAAG AGCAAGGAGG AGGCCAGCA 360
CGTCCAGGGA GTACTGGCCC AGCTCCTGAG GCGGGAGGCA GCCCTGACGG CGAGGATGAG 420
CAAGTTCCTG ATTGGGCTCC AGCGAGGAA GGGCAAGTGC CTGGACCCCTA GTCTGCCGCT 480
GAAGGGCTTC AGCTGGGTGG GCGGGGGGGA GGACACGCTT TACTCTAACT GGCACAAGGA 540
GCTCCGGAAC TCGTGCATCT CCAAGCGCTG TGTGTCTCTG CTGCTGGACC TGTCCAGGCC 600
GCTCCTTCCC AACCGCTGCG CCAAGTGGTG TGAGGGCCCC TGTGGGAGCC CAGGCTCCCC 660
CGGAAGTAAC ATTGAGGGCT TCGTGTGCAA GTTCAGCTTC AAAGGCATGT GCCGCGCTCT 720
GGCCCTGGGG GGGCCAGGTC AGGTGACCTA CACCACCCCTT TCCAGACCA CCAGTTCCTC 780
CTTGGAGGCT TAGCCTTTTG CCTCTGCGGC CAATGTAGCC TGTGGGGAAG GTGACAAGGA 840
CGAGACTCAG AGTCATTATT TCCTGTGCAA GGAGAGGCC CCGATGTGT TCGACTGGGG 900
CAGCTGGGCG CCCCTCTGTG TCAGCCCCAA GTATGGCTGC AACTTCAACA ATGGGGCTG 960
CCACCAAGAC TGTCTTGAAG GGGGGGATGG CTCTTCTCTG TCGCGCTGCC GACCAGGATT 1020
CCGGCTGCTG GATGACTTGG TGACCTGTGC CTCTCGAAAC CCTTGCAGCT CCAGCCCATG 1080
TCGTGGGGGG GCCAGGTGGG TCCTGGGACC CCATGGGAAA AACTACACGT GCGCGTCCCC 1140
CCAAGGGTAC CAGCTGGACT CGAGTCAGCT GGACTGTGTG GACGTGGATG AATGCCAGGA 1200
CTCCCTCTGT GCCCAGGAGT GTGTCAACAC CCCTGGGGGC TTCCGCTCGG AATGCTGGGT 1260
TGGCTATGAG CGGGGCGGTG CTGGAGAGGG GGCCTGTCTG GATGTGGATG AGTGTCTCT 1320
GGGTGCTGCG CCTTGCGCCG AGGGCTGCAC CAACACAGAT GGTCTATTTC ACTGCTCCTG 1380
TGAGGAGGGC TAGCTCTTGG CCGGGGAGGA CGGGAAGTGC GTCTCTTGA CCATGGGGCC 1440
TGTGGGCCCC GGGGGCCCCC TCTGCGACAG CTTGTGCTTC AACACACAG GGTCTCTCA 1500
CTGTGGCTCG CTGCCAGGCT GGGTGTCTGC CCCAAATGGG GTCTCTTGA CCATGGGGCC 1560
TGTGTCTCTG GACCAACCAT CTGGGCCCCC CGATGAGGAG GACAAAGGAG AGAAGGAAG 1620
GAGCACCGTG CCGCGCGCTG CAACAGCCAG TCCACAAAGG GGGCCCGAGG GCACCCCCAA 1680
GGCTACACCC ACCACAAGTA GACCTTCGCT GTCATCTGAC GCGCCCATCA CATCTGCCCC 1740
ACTCAAGATG CTGGCCCCCA GTGGGTCTCT AGGCGTCTGG AGGGAGCCCA GCATCCATCA 1800
CGCCACAGCT GCCTCTGGCG CCCAGGAGCC TGCAAGTGGG GACTCTCTCG TGGCCACACA 1860
AAACAACGAT GGCACGTAGG GGCAAAAGCT GCTTTTATTC TACATCTAG GCACCGTGGT 1920
GGCCATCTTA CTCTGCTGCG CCCTGGCTCT GGGGCTACTG GTCTATCGCA AGCGGAGAGT 1980
GAGAGGGGAG GAGAAGAAGG AGAAGAAGCC CCAGAATGGG GCAGACAGTT ACTCCTGGGT 2040
TCCAGAGCGA GCTGAGAGCA GGGCCATGGA GAACCAATAC AGTCCGACAC CTGGGACAGA 2100
CTGCTGAAAG TGAGTGGGCC CTAGAGACAC TAGATCACC AGCCACCATC CTCAGAGCTT 2160
TGAACCTCCC ATTCCAAAGG GGCACCCACA TTTTCTTGA AGACTGGACT GGAATCTTAG 2220
CAAACTATTT TAAGTCTCTC CTCTAAAGGC CCCTTGGAAC ATGCAGGTAT TTTCTACGGG 2280
TGTTTGATGT TCTTGAAGTG GAAGCTGTGT GTTGGCGTGC CACGGTGGGG ATTTCTGTAC 2340
TCTATAATGA TTGTTACTCC CCCTCCCTTT TCAAATTTCA ATGTGACCAA TTCCGGATCA 2400
GGGTGTGAGG AGGCTGGGGC TAAGGGGCTC CCTGAATAT CTCTCTGCT CACTTCCACC 2460
ATCTAAGAGG AAAAGGTGAG TTGCTCATGC TGATTAGGAT TGAATGATT TGTCTCTCT 2520
CCTAGGATGA AAATCAATTC AATTAATTAT TCAATTAGGT AAGAAGATCT GGTTTTGGG 2580
TCAAAGGGAA CATGTCGGA CTGGAACAT TTCTTTACAT TTGCATTCTT CCATTTCGCC 2640
AGCACAAAGT TTGCTAAATG TGATACTGTT GACATCTCTC AGAATGSCCA GAAGTGAAT 2700
TAACTCTTA GGTGGCAAGG AGGCAGGAAG TGCCCTCTTA GTTCTTACAT TTCTAATAGT 2760
CTTGGGTTTA TTTGCAAGG AAGCTTGAAA AATATGAGAA AAGTGTCTTG AAGTGCATTA 2820
CAGGTGTTTG TGAAGTCACA TAATCTACGG GGTAGGGGCG AGAGAGGCCA GGGATTGTT 2880
CACAGATACT TGAATTAATT CATCCAAATG TACTGAGGTT ACCACACAT TGAATACGGA 2940
TGTGATCAAC ACTAACAGG AAACAAATTC AAGGACAACT TGTCTTTGAG CCAGGCGCAGG 3000
CCTCAGACAC CTGCTCTGTG GCGCCGCTCT CACTTCAATC TGCCCGGAAT GCCAGTGCTC 3060
CGAGCTCAGA CAGAGGAAGC CCTGCAGAAA GTTCCATCAG GCTGTTTCTT AAAGGATGTG 3120
TGAACGGGAG ATGATGCACT GTGTTTGAAG AGTGTCTCAT TTAAGCAATT TTAGCACAGT 3180
TCATAGTCCA CAGTTGATGC AGCATCTGGA GATTTTAAAT CCTGAAGTGT GGGTGGCGCA 3240
CACACCAAGT AGGGAGCTAG TCAGGCGATT TGCTTAAAGG ACTTTTGTTC TCTGTCTCTT 3300
TTCTTAAAAA TTGGGGTAA GGAGGGGAAG AAGAGGGAAA GAGATGACTA ACTAAATCA 3360
TTTTTACAGC AAAAATGCT CAAAGCCATT TAAATTATAT CCTCATTTTA AAAGTTACAT 3420
TTGCAATAT TTCTCCCTAT GATAATGCAG TCGATAGTGT GCACTCTTTT TCTCTCTCTC 3480
TCTCTCTCAC ACACACACAC ACACACACAC ACACACACAC AGAGACAAGG CACCATCTCTG 3540
CCTGGGGCAC TGGAACACAT TCCTGGGGGT CACCGATGGT CAGAGTCACT AGAAGTTACC 3600
TGAGTATCTC TGGGAGGCCT CATGTCTCCT GTGGGCTTTT TACCACCACT GTGCAGGAGA 3660
ACAGACAGAG GAAATGTGTC TCCTTCCAAG GCCCCAAAGC CTCAGAGAAA GGGTGTCTCT 3720
GGTTTTGCCT TAGCAATGCA TCGTCTCTG AGGTGACACT CTGGAGTGGT TGAAGGGCCA 3780
CAAGGTGCAG GGTTAATACT CTGCGCAGTT TTGAAATATA GATGCTATG TTAGATTGT 3840
TTTTAATAGA AAATCAAAGG GGCAGGGGAA GTGAAAGGAA AGATGGAGGT TTTGTGCGGC 3900
TOGATGGGGC ATTTGGAATC TCTTTTAAAT GTCATCTCAT GGTCTCCAGT TTTCAAGTTG 3960
AACTCTGGTG TTTAACAATT AAGGGAGACA AAGGCTGTGT CCATTGGGCA AAATCTCCTT 4020
GGCCACGAGA CTCTAGGTGA TGTGTGAAGC TGGGCACTCT GTGGTGTGGA GAGCAGCCAT 4080
CTGTCTGGCC ATTCAGAGGA TTCTAAAGAC ATGGCTGGAT GCGCTGTGTA CCAACATCAG 4140
CACTTAAATA AATGCAAAAT CAACATTCTC COCTCTGGGC CTTGAAATC CTGCGCTTA 4200
TCAITTTGGG TGAAGGAGAC ATTTCTGTCC TTGGCTTCCC ACAGCCCCAA CGCAGTCTGT 4260
GTATGATTCG TGGATCCAA CGAGCCCTCC TATTTTCACA GTGTTCTGAT TGCTCTCACA 4320
GCCAAGGCC TTGTTTGGGG GATCTTGGGG CTACAGTAGT GGGTAAACAA ATGCCACCG 4380
GAACCCCTCT GTGGAACCCA CAAGGGGAGA AATGGGTGAT AAAGAATCCA GTTCTCTCAA 4440
ACCTTCCCTG GCAGGCTGGG TCCCTCTCCT GCTGGGTGCT GCTTCTCTT GCACACCACT 4500
CCCAACACGG GGGGAGAGCC AGCAACCCAA CCAGACAGCT CAGGTGTGTC ATCTGATGGA 4560
AACCAGTGGG CTCACACAGC TGCTTTATTC TCCTGTTTAT TTTTGTGTTT ACTTTGAAGC 4620
ATGGAATTC TTGTTTGGGG GATCTTGGGG CTACAGTAGT GGGTAAACAA ATGCCACCG 4680
GCCAAGAGGC CATTAAACAA TCGTCTCTGT CTGAGGGGGC CCAGCTTGC TCGGGCTGTG 4740
CACAGTGGGG AATCCAAGGG TCACAGTATG GGGAGAGGTG CACCTTGCCA CTTGCTAACT 4800

```

TCTGCTAGA CACAGTCTTT CTGCCCAGGT GACCTGTTC GCAGCAGAAC AAGCCAGGGC 4860  
 CATGGGAGAG GGGGAGTTT TCACITGGAG ATGGACACCA AGACAATGAA GATTTGTGTG 4920  
 CCAATATAGT CAATATTTCT GGGAGACTCT TGGAAAAAAC TGAATATATT CAGGACCAAC 4980  
 5 TCTCTCCCTC CCCTCATCCC ACATCTCAAA GCAGACAATG TAAAGAGAGA ACATCTCACA 5040  
 CACCCAGCTC GCCATGCCTA CTCATCTCTG AATTTAGGT GCCATCAGT CTCTTTCTTT 5100  
 CTCTTTTGTG ATTTAGAAA GGATGCAGGA GGACAATTCC CACAGATAAT CTGAGGAATG 5160  
 CAGAAAAACC AGGGCAGGAC AGTTATCGAC AATGCATTAG AACITGGTGA GCATCCTCTG 5220  
 TAGAGGAGCT CCACCCCTGC TCAACAGCTT GGCTTCCAGG CAAGACCAAC CACATCTGGT 5280  
 10 CTCTGCCCTC GGTGGCCAC ACACCTAAGC GTCATCGTCA TTGCCATAGC ATCATGATGC 5340  
 AACACATCTA CGTGTAGCAC TACGACGTTA TGTTTGGGTA ATGTGGGGAT GAACCTGCATG 5400  
 AGGCTCTGAT TAAGATGTG GGGAGTGGG CTGCGGTAC TGTCGGCTT GCAAGGCCAC 5460  
 CTGGAGGCTT GTCTGTAGC CAGTGGTGA GGAGCAAGGC TTCAGGAAGG GCCAGCCACA 5520  
 TGCCATCTTC CCTGCGATCA GGCAAAAAG TGAATATAA AAGTCAAAAC TTTATATGCA 5580  
 15 TGTGTTATGT CCATTTTGA GGATGAACCT AGTTTAAAG AATTTTTTT TCTCTCAAG 5640  
 TTGCTTTGTC TTTTCCATCC TCATCACAAG CCCTTGTGTT AGTGTCTTAT CCCTGAGCAA 5700  
 TCTTTGATG GATGAGATG ATCATTAGT ACITTTGTTT CAACCTTTAT TCTCTTAAAT 5760  
 ATTTCTGTGA AAACTAGAG AACAGAGAT AGATTGACA AAAAAAATT GAATTAATAA 5820  
 TAACACAGTC TTTTAAAC TAACATAGGA AAGCCTTTC TATTATTCT CTCTTAGCT 5880  
 20 TCTCAATTG TAAATCAGG AAAACAGGAA AACACAGCT TCTAGCAGT GCAAAATGTT 5940  
 TTAATGCCCC CTACATATT CCATCACCTT GAACAATAGC TTAGCTTGG GAATCTGAGA 6000  
 TATGATCCCA GAACACATCT GTCTCTACT CGCTGCAAA ACCCATGGTT TAAATCTATA 6060  
 TGGTTGTGTC ATTTCTCAA CTAAAAATAG AGATGATAAT CCGAATTCTC CATATATTCA 6120  
 CTAATCAAAG ACATATTTT CATACTAGAT TCCTGAGACA AATACTCACT GAAGGGCTTG 6180  
 25 TTTAAATAA AATTGTGTT TGTCTGTTT TGTAGATAA TGCCCTTCTA TTTTAGGTAG 6240  
 AAGCTCTGGA ATCCCTTAT TGTGCTGTT CTCTTATCTG CAAGGTGGCA AGCAGTCTCT 6300  
 TTAGCAGAT TTTGCCATG ATTCCTCTGA GCTGAAGTTC TTGTCATAGA TTTGGCTTAA 6360  
 GCTTGAATTA GATCCTGCA AAGGCTTGT CTGTGATGTC AGATGTAAT GTAATGTCA 6420  
 30 GTAATCACT CATGAATGT AAATGAGAT GTAAGTATT TTAATGTGT GTAATTCAAA 6480  
 TTTGTTTGC TAATCTGGA ATTACAAGAT TTCTATGCAG GATTACCTT CATCTGTGC 6540  
 ATGTTTCCCA AACTGTGAGG AGGGAAGGCT CAGAGATCGA GCTTCTCTC TGAGTTCTAA 6600  
 CAAAATGTT CTTTAGGGT CAGCCTTTAG GAAGGTGCAG CTTTGTGTC CTTTGAAGCT 6660  
 TCTGTTATGT GCCTATCCTA ATAACTCTT AAACACATT

Seq ID NO: 345 Protein sequence  
Protein Accession #: NP\_036204

1 11 21 31 41 51  
 MATSMGLLLL LLLLLTPGA GTGADTEAVV CVGTACYTAH SGKLSAAEAQ NHCNQNGNL 60  
 40 ATVKSKEBAQ HVQRLAQLL RREAALTARM SKFWIGLQRE KGKCLDPSLP LKQPSWVGG 120  
 EDTPYSNHKK ELRNGSCISK CVSLLDLSQ PLLPNRLPKW SEGPGSGSPS PGSNIEGFVC 180  
 KFSFKMCRP LALGPGQVTT YTPPTQTS SLEAVPFASA ANVACGEGDK DETQSHYFLC 240  
 KEKAPDVPDW GSSGPLCVSP KYGCNPNNGG CHQDCFEGBD GSFLQGRPG FRLLDLDVTC 300  
 45 ASRNPCCSSP CRGGATCVLG PHGKNYTCRC PQGYLDSSQ LDCVDVDEQ DSPCAQECVN 360  
 TPGSPRCEBV VGYEPGPGOE GACQDVDECA LGRSPCAQGC TNDGSPHCS CEEGYVLAGE 420  
 DGTQCDVDWE VYGPGLDCE SLCFNTQSSP HCGCLPGWVL APNGVSCVMG PVSIGPPSPG 480  
 PDDEKDGKKE GSTVPRAATA SPTRPGPTP KATPTTSRPS LSSDAPITSA PLKMLAPSGS 540  
 50 SGVWREPSIH HATAASGPQE PAGGDSVAT QNNDGTDGQK LLLFYILGTV VAILLLALA 600  
 LGLLVYRKR AKREKKEKK PQNAADSYW VPERAESRAM ENQYSPTPGT DC

Seq ID NO: 346 DNA sequence  
Nucleic Acid Accession #: Z31560  
Coding sequence: <1-966

1 11 21 31 41 51  
 CACAGCGCCC GCATGTACAA CATGATGGAG ACAGGAGCTGA AGCCGCGGGG CCCGAGCAA 60  
 ACTTCGGGGG GCGGCGGCGG CAACTCCACC GCGGCGGCGG CCGGCGGCAA CCAGAAAAAC 120  
 60 AGCCCGGACC GCGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CGGGCAGCGG 180  
 CGCAAGATGG CCCAGGAGAA CCCCAAGATG CACAACCTCG AGATCAGCAA GCGCCTGGCG 240  
 GCGGAGTGA AACTTTTGTG GGAGACGGAG AAGCGGCGGT TCATCGACGA GGCTAAGCGG 300  
 CTGCGAGCGC TGCACATGAA GGAGCACCGG GATTATAAAT ACCGGCCCGG GCGGAAAAAC 360  
 AAGACGCTCA TGAAGAAGGA TAAGTACAGC CTGCGCGGCG GGCTGCTGGC CCCGCGCGGC 420  
 65 AATAGCATGG CGAGCGGGGT CCGGTGGGCG GCCCGCTGCG GCGCGGCGGT GAACGAGCGC 480  
 ATGGACAGTT ACGCGCATAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGAGGAC 540  
 CAGCTGGGCT ACCCGCAGCA CCGGGGCTCT AATGCGCAGC GCGCAGCGCA GATGAGCGCC 600  
 ATGACCGCT ACAGCTGAG CGCCCTGCAG TACAACCTCA TGACAGCTC GCAGACCTAC 660  
 ATGAACGGCT CGCCACCTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720  
 70 CTGGCTCCA TGGGTTCGGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCCT TGTGTTACC 780  
 TCTTCTCCCT ACTCCAGGGC GCCCTGCCAG GCGGGGACCC TCGGGGACAT GATCAGCATG 840  
 TATCTCCCGG GCGCGAGGT GCGGGAACCC GCGGCCCCCA CGAGACTTCA CATGTCCCAG 900  
 CACTACCAGA GCGGCGCGGT GCGGCGCAGC GCCATTAAAG GCACACTGCC CCTCTCACAC 960  
 ATGTAGGGG CCGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAGAAAAA ACGAGGGAAA 1020  
 75 TGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCCGG TACGCTCAAA 1080  
 AAAAA

Seq ID NO: 347 Protein sequence  
Protein Accession #: CAA83435

1 11 21 31 41 51  
 HSARMYNME TELKPPGPQQ TSGGGGGNST AAAAGGNQKN SPDRVKRPMN AFMVNSRGQR 60  
 80 RMAQENPKM HNSBISKRLG AEWKLLSETE KRPFIDEAKR LRALHMKHP DYKYRPRRET 120  
 KTLMKDKYT LPPGLAPGG NSMAGVGVG AGLGAGVQNR MDSYAHNGW SNGYSMMQD 180  
 85 QLQYPQHPGL NAHGAQMMP MERYDVSAIQ YNSMTSSQTY MNGSPYYSMS YSQGTFPMA 240  
 LGSMGSPVKS EASSSPVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPPEP AAPSRLLHMSQ 300  
 HYQSPVPGT AINGTLPLSH M

Seq ID NO: 348 DNA sequence  
Nucleotide Accession #: NM\_002638  
Coding sequence: 120-473

```
1      11      21      31      41      51
|      |      |      |      |      |
CAATACAGCT AAGGAATTAT CCCTTGTAAG TACCACAGAC CGCCCTGGA GCCAGGCCAA 60
GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCCTCAT CGCTGGGACG CTGGTCTTAG 180
AGGCAGCTGT CAGCGGAGTT CCTGTTAAAG GTCAGACAC TGTCAAAGGC CGTGTTCAT 240
TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300
CGCAAGAGCC AGTCAAAGGT CCAAGTCTCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
TCCGGTGGCG CATGTTGAAT CCCCTAACC GCTGCTTGA AGATACTGAC TGCCCAGGAA 420
TCAAGAAAGT CTGTGAAGGC TCTTGCGGGA TGGCCTGTT CGTTCGCCAG TGAAGGGAGC 480
CGGTCTTGGC TGCACCTGTG CCGTCCCCAG AGCTACAGGC CCCATCTGGT CCTAAGTCCC 540
TGCTGCCCTT CCCTTCCCA CACTGTCCAT TCTTCTCCC ATTCAAGATG CCCACGGCTG 600
GAGCTGCCCT TCTCATCCAC TTTCCAATAA A
```

Seq ID NO: 349 Protein sequence:  
Protein Accession #: NP\_002629

```
1      11      21      31      41      51
|      |      |      |      |      |
MRASSFLIVV VPLIAGTLVL EAAVTGVFVK GQDTVKGVRP FNGQDPVKQ VSVKQDKVK 60
AQEPVKGPSV TKPGSCPIIL IRCAMLNFPN RCLKDTDCPG IKKCEGSCG MACFVPQ
```

Seq ID NO: 350 DNA sequence  
Nucleic Acid Accession #: NM\_007183  
Coding sequence: 75-2468

```
1      11      21      31      41      51
|      |      |      |      |      |
GAATTCGGGA CAGGACGTGA AGATAGTTGG GTTTGGAGGC GGCCGCCAGG CCCAGGCCCG 60
GTGGAACTGC CGCCTATGAG GACGGTAAT TCCTGCTGTC GGCCCTGCAG CCTGAGGCCG 120
CGGTGTGCTC CTGCGCGCTG CCCTCTGACC TGCAGCTGGA CGCCCGGGGC GCCAGGGGGC 180
CGGAGCGCGA GCGCTGCGG GAGCCCGCG TCCAGGAGCA GGTCCGCGCC CGCTCTTTGC 240
AGCTGGGACA GCAGCCGCGG CACAACGGGG CCGCTGAGCC CGAGCCTGAG GCCGAGACTG 300
CCAGAGGCAC ATCAGGGGGG CAGTACCACA CCCTGCAGGC TGGCTTCAGC TCTCGCTCTC 360
AGGGCTGAGT TGGGACAAAG ACCTCGGGCT TCCGGCCCAT CGCCAAAGCG GCCTACAGCC 420
CAGCCTCTCT GTCTCTCCGC TCCGCGGTGG ATCTGAGCTG CAGTCCGAGG CTGAGTTTCA 480
CCCACAATGG GGGCAGCGCC TTTGGGGCGG CTGGGTACGG GGGTGCCTCG CCCACCCCTC 540
CCATGCCAC CAGGCCCGTG TCCTTCCATG AGCGCGGTGG GGTGGGAGC CGGGCCGACT 600
ATGACACACT TCCTCTGCGC TCGCTGCGGC TGGGGCCCGG GGGCTGGAG GACCCGTACA 660
GCTGTGCTTC TGAGCAGCTG GAGCCCGCGG CCACTCCAC GTACAGGGCC TTTGCTTACG 720
AGCGCCAGCG CAGCTCCAGC TCCAGCGCGG CAGGGGGGCT GGACTGGCCC GAGGCCACTG 780
AGGTTTCCCC GAGCGGAGCC ATCCGTGCCC CTGCGGTGCG GACCTGTCAG CGATTCCAGA 840
GCAGCCACCG GAGCGCGGGG GTAGGCGGGG CAGTGCCGGG GGCGCTCTG GAGCCAGTGG 900
CTCGAGCGCC ATCTGTGCGC AGCCTCAGCC TCAGCCTGGC TGACTCGGGC CACTGCGCG 960
ACGTGCATGG GTTCAACAGC TACGGTAGCC ACGGAACCTT GCAGAGACTC AGCAGCGGTT 1020
TTGATGACAT TGACCTGCCC TCAGCAGTCA AGTAACCTCAT GGCTTCAGAC CCCAACCTGC 1080
AGGTGCTGGG AGCGGCTTAC ATCCAGCACA AGTGCTACAG CGATGCAGCC GCCAAGAAGC 1140
AGGCCCGCAG CCTTCAGGCC GTGCCTAGGC TGGTGAAGCT CTTCAACCA GCGAACAGG 1200
AAGTGACCGC CCATGCCACA GGTGCCATGC GCAACCTCAT CTAAGACAAC GTGACAACA 1260
AGCTGGCCCT GGTGGAGGAG AACGGGATCT TCGAGCTGCT GCGGACACTG CGGGAGCAGG 1320
ATGATGAGCT TCGCAAAAT GTCAAGGGA TCCTGTGGAA CCTTTCATCC AGCGACCACC 1380
TGAAGGACCG CTTGGCCACA GACAAGCTGG AGCAGCTCAC GGACCTGGTG TTGAGCCCCC 1440
TGTGCGGGGC TGGGGGTCCC CCCTCATCC AGCAGAACGC CTCGAGGCGG GAGATCTTCT 1500
ACAACGCCAC CGCTTCTCT AGGAACCTCA GCTCAGCCTC TCAGGCCACT CGCCAGAAGA 1560
TGCGGGAGTG CACCGGGCTG GTGGACGCCC TGGTCACTTC TATCAACCA GCCCTGAGCG 1620
CGGGCAATG CGAGGACAAG AGCGTGGAGA ACGCGGTGTG CTCTCTGGG AACCTGTCTT 1680
ACCGCTCTA CGACGAGATG CCGCGTCCG CGCTGCAGCG GCTGGAGGGT CGCGGCCGCA 1740
GGGACCTGGC GGGGGCGCG CCGGAGAGG TCGTGGGCTG CTTACGCGG CAGAGCCGGC 1800
GGCTGCGGGA GCTGCCCTCT GCGCGGATG CGCTCACCTT CCGGAGGTG TCCAAGGAOC 1860
CCAAGGGCCT CGAGTGGCTG TGGAGCCCC AGATCGTGGG GCTGTACAAC CGGTCTGTGC 1920
AGCGCTGCGA GCTCAACCGG CACACGACGG AGCGGGCGCG CGGGGCGCTG CAGAACATCA 1980
CGGCAGGCGA CCGCAGGTGG GCGGGGTGTC TGAGCCGCTT GGCCCTGGAG CAGGAGCGTA 2040
TTCTGAACCC CTGTGTAGAC CGTGTGAGG CCGCGGACCA CCACCAGCTG CGCTCACTGA 2100
CTGGCTCAT CCGAAACCTG TCTCGGAACG CTAGGAACAA GGACGAGATG TCCAAGGAGG 2160
TGGTGAAGCA CTGTATCGAG AAGCTGCCAG GCAGCGTGGG TGAGAAGTGG CCCCAGCCG 2220
AGGTGCTGGT CAACATCATA GCTGTGCTCA ACAACCTGGT GGTGGCCAGC CCTATCGCTG 2280
CCGAGAGCT GCTGTATTTT GACGGACTCC GAAAGCTCAT CTTTCATCAAG AAGAAGCGGG 2340
ACAGCCCCGA CAGTGAGGAG TCCTCCCGGG CAGCATCCAG CCTCTGGGCC AACCTGTGGC 2400
AGTACAACAA GCTCCACCGT GACTTTCGGG CGAAGGGCTA TCGGAAGGAG GACTTCTTGG 2460
GCCCATAGGT GAAGCTTCTT GGAGGAGAAG GTGAGTGGC CCAGGTCCA AGGGACAGAC 2520
TCAGTCCAG GCTGCTTGGC AGCCAGGCTT GGAGGAGAAG GCTAATGACG GAGGGGCCCC 2580
TCGCTGGGGC CCTGTGTGCT ATCTTTGAGG GTCTGGGCC ACCAGGAGGG GCAGGCTCTT 2640
ATAGTGGGG ACTTGGCTTC CGCAGGGCAG GGGGTGGGG AGGGCTCAAG GCTGCTCTGG 2700
TGTATGGGGT GGTGACCCAG TCACATTGGC AGAGGTGGGG GTTGGCTGTG GCCTGGCAGT 2760
ATCTGGGAT AGCCAGCACT GGGAAATAAG ATGGCCATGA ACAGTCACAA AAAAAAAAAA 2820
AAAAGGAATT C
```

Seq ID NO: 351 Protein sequence  
Protein Accession #: NP\_009114.1

```
1      11      21      31      41      51
```

MQDGNFLLSA LQPEAGVCSL ALPSDLQLDR RGAEGPEAER LRAARVQEQV RARLLQLGQQ 60  
PRHNGAAEPE PEAEATAGTS RQYHTLQAG FSSRSQGLSG DKTSGFRPIA KPAYS PASWS 120  
SRSAVLDLSCS RRLSSAHNGG SAFGAAGYGG AQPTPPMPTR PVSFHERGGV GSRADYDTLS 180  
LRSRLGPGG LDRYSLVSE QLEPAATSTY RAFAYERQAS SSSSRAGGLD WPEATEVSFS 240  
RTIRAPAVRT LQRFQSSHRS RGVGGA VPGA VLEPVARAPS VRSLSLSLAD SGHLPDVHGF 300  
NSYGSRTTLQ RLSSGDDID LPSAVKYLMA SDPNLQVLGA AYIQHKCYSD AAKKQARS 360  
QAVPRILVKLF NHANQEVQRH ATGAMRNLIY DNADNKLALV EENGIFELLR TLREQDDELR 420  
KNVTGLLWNL SSSDHLKDRLL ARDTLEQLTD LVLSPLSGAG GPPLIQONAS EAEIFYNATG 480  
FLRNLSSASQ ATRQKMRECH GLVDALVTSI NHALDAGKCE DKSVENAVCV LRNLGYRLYD 540  
EMPPSALQRL EGRGRRLDLAG APPGEVVGCF TPQSRRLREL PLAADALTFA EVSKDPKGLE 600  
WLNSPQIVGL YNRLLQRCEL NRHTTEAAAG ALQNTAGDR RNAGVLSRLA LEQERILNPL 660  
LDRVRTADEH QLRSLTGLIR NLSRNARNKO EMSTKVVSHL IEKLPGSVGE KSPPAEVLVN 720  
IIAVLNNLVV ASPIAADLL YFDGLRKLIF IKKRDSPDS EKSSRAASSL LANLNQYNKL 780  
HRDFRAKGYR KEDFLGP

Seq ID NO: 352 DNA sequence

Nucleic Acid Accession #: M31469

Coding sequence: 1-651

1 11 21 31 41 51  
ATGGCTGCGC AGGAGAGGCC CCAGGTCCAG TTCAAACCTTG TATTGGTTGG TGATGGTGGT 60  
ACTGGAAAAA CGACCTTCGT GAAACGTCAT TTGACTGGTG AATTGAGAA GAAGTATGTA 120  
GCCACCTTGG GTGTTGAGGT TCATCCCTTA GTGTCCACA CCAACAGAGG ACCTATTAG 180  
TTCAATGTAT GGGACACAGC CGGCCAGGAG AAATTCGGTG GACTGAGAGA TGGCTATTAT 240  
ATCCAAGCCC AGTGTGCCAT CATAATGTTT GATGTAACT CGAGAGTTAC TTACAAGAAT 300  
GTGCCTAAT CGCATAGAGA TCTGGTACGA GTGTGTGAAA ACATCCCAT TGTGTGTGT 360  
GGCAACAAAG TGGATATTAA GGACAGGAAA GTGAAGCGGA AATCCATTGT CTTCACCGA 420  
AAGAAGAATC TGCAGTACTA CGACATTTCT GCCAAAAGTA ACTACAACCT TGAAAAGCCC 480  
TTCTCTGCGC TTGCTAGGAA GCTCATTTGA GACCTTAAT TGAATTTGT TGCCATGCCT 540  
GCTCTCGCCC CACCAGAGT TGTCTATGAC CCAGCTTTGG CAGCACAGTA TGACGACGAC 600  
TTAGAGGTTG CTCAGACAACT TGCTCTCCCG GATGAGGATG ATGACCTGTG A

Seq ID NO: 353 Protein sequence

Protein Accession #: AAA36546

1 11 21 31 41 51  
MAAQGEPOVQ FKLVLVGDGG TGKTFVVRH LTGEPEKKYV ATLGEVEHPL VPHNTRGPIK 60  
FNVMVDTAGS KFGGLRDYGY IQAQCAIIMF DVTSRVTYKN VFNWHRDLVR VCENIPIVLC 120  
KNKVDIKDRK VKAKSIVFHR KKNLQYDYS AKSNYNFEKP FLWLARKLIG DPNLEFVAMP 180  
ALAPPEVVMV PALAAQYEHF LEVAQTALP DEDDDL

Seq ID NO: 354 DNA sequence

Nucleic Acid Accession #: NM\_002820

Coding sequence: 304-831

1 11 21 31 41 51  
COGGTTGCGA AAGAAGCTGA CTTAGAGGG GGAACCTTC TTCTTTTAGG AGGCGGTTAG 60  
CCCTGTTCCA CGAACCCAGG AGAAGTCTG GCCAGATTAA TTAGACATTG CTATGGGAGA 120  
CGTGTAACCA CACTACTTAT CATTGATGCA TATATAAAC CATTATTATT TCGCTATTAT 180  
TTAGAGGAAA GCGCCTCTGA TTTGTTTCTT TTTTCCCTTT TTGCTCTTTC TGGCTGTGTG 240  
GTTTGGAGAA AGCAGAGTTG GAGTAGCGGG TTGCTAAATA AGTCCCGAGC GCGAGCGGAG 300  
ACGATGTCAGC GGAAGCTGGT TCAGCAGTGG AGCGTCCGGG TGTCTCTGCT GAGCTACGCG 360  
GTGCTCTCTC GCGGCGCTCG GGTGGAGGGT CTCAGCCGCC GCCTCAAAAG AGCTGTGTCT 420  
GAACATCAGC TCCTCCATGA CAAGGGGAAG TCCATCCAAG ATTTACGGCG ACGATTCTTC 480  
CTTCACTATC TGATCGCAGA AATCCACACA GCTGAAATCA GAGCTACCTC GGAGGTGTCC 540  
CCTAACTCCA AGCCCTCTCC CAACACAAAG AACCACCCCG TCCGATTGGT GTCTGATGAT 600  
GAGGGCAGAT ACCTAATCTA GGAACTAAC AAGGTGGAGA CGTACAAAGA GCAGCCGCTC 660  
AAGACACCTG GGAAGAAAAA GAAAGGCAAG CCGGGGAAAC GCAAGGAGCA GGAAGAGAAA 720  
AAACGGGAAA CTCGCTCTGC CTGGTTAGAC TCTGGAGTGA CTGGGAGTGG GCTAGAAGGG 780  
GACCACCTGT CTGACACCTC CACAACGTCG CTGAGCTCG ATTACAGGTA ACAGGCTTCT 840  
CTGGCCCGTA GCTCAGCGG GGTGCTCTCA GCTGGGTTT GGAGCTCCC TTCTGCTGTG 900  
GCTTGGACAA ACCTAGAATT TTCTCCCTTT ATGTATCTCT ATCGATTGTG TAGCAATTGA 960  
CAGAGAATAA CTCAGAAATAT TGTCTGCCTT AAAGCAGTAC CCCCCTACCA CACACACCCC 1020  
TGTCTCTCAG CACCATAGAG AGGCGCTAGA GCCCATTCCT CTCTCTCCAC GGTCAACCAA 1080  
CATCAATCCT TTACCATCTT ACCAAATAAT TTCAATATCA AGCTTCAGAA GCTAGTGACC 1140  
ATCTTCATAA TTTGCTGGAG AAGTGTATTT CTCCCTCTA CTCTCACACC TGGGCAAACT 1200  
TTCTTCAGTG TTTTTCATTT CTACGTTCTT TTCACCTCAA GGGAGAAATAT AGAAGCATTT 1260  
GATATTATCT ACAACACCTG CAGAACAGCA TCATGTCTA AAGCATTCG AGCCATTCTC 1320  
ACTTTTATTT TAATTAAATG TATTTAATTA AATCTCAAAT TTATTTTAAT GTAAAGAACT 1380  
TAAATTATGT TTTAAACACA TGCCCTTAAAT TTGTTTAAAT AAATTTAACT CTGTTTCTTA 1440  
CCAGCTCATA CAAATAAATG GGTTCCTGAA AATGTTTAA TATTAACCTA CAAGGATATA 1500  
GGTTTTCCTC ATGTATCTTT TTGTTTATTG GCAAGATGAA ATAATTTTTC TAGGGTAATG 1560  
CGTAGGAAA AATAAAACTT CACATTTAAA AAAA

Seq ID NO: 355 Protein sequence

Protein Accession #: NM\_002820

1 11 21 31 41 51  
MQRRVLQQWS VAVFLSYAV PSCGRSVEGL SRRLKRAVSE HQLLHDKGKS IQDLRRRFFL 60  
HHLIAIEHTA EIRATSEVSP NSKPSPNTKN HPVRFSGSDE GRYLQTETNK VETYKEQPLK 120

TPGKKKRGKP GRRKEQKKK RRTRSAWLDS GVTGSGLEGD HLSDTSTTSL ELDSR

Seq ID NO: 356 DNA sequence

Nucleic Acid Accession #: NM\_017522

Coding sequence: 1-2100

5  
10  
15  
20  
25  
30  
35  
40  
45  
50

1 11 21 31 41 51  
ATGGGCTCC CCGAGCCGGG CCTCTCCGG CTTCGCGCG TGCTGCTGCT GCTGCTGCTG 60  
CTGCTGCTGC TGCGGCTCCA GCATCTTGCG GCGGCAGCG CTGATCCGCT GCTCGGCGGC 120  
CAAGGGCCGG CCAAGGAGTG CGAAAAGGAC CAATTCCAGT GCCGGAACGA GCGCTGCATC 180  
CCCTCTGTGT GGAGATGCGA CGAGGACGAT GACTGCTTAG ACCACAGCGA CGAGGACGAC 240  
TGCCCAAGA AGACCTGTGC AGACAGTGAC TTCACCTGTG ACAAAGGCCA CTGCATCCAC 300  
GAACGGTGA AGTGTGACCG CGAGGAGGAG TGTCTGATG GCTCCGATGA GTCGAGGCC 360  
ACTTGACCA AGCAGGTGTG TCCTGCAGAG AAGCTGAGCT GTGAGCCAC CAGCCACAAG 420  
TGTGTACTG CTCTGTGGCG CTGCGAGCGG GAGAAGGACT GCGAGGGTGG AGCGGATGAG 480  
GCGCGTGTG CTACCTCACT GGGCACTGCG CGTGGGGAG AGTTCAGTG TGGGGATGGG 540  
ACATGTGTCC TTGCAATCAA GCACTGCAAC CAGGAGCAG ACTGTCCAGA TGGGAGTGAT 600  
GAAGCTGGCT GCCTACAGGG GCTGAACGAG TGTCTGCACA ACAATGGCGG CTGCTCACAC 660  
ATCTGCACTG ACCTCAAGAT TGGCTTTGAA TGCACTGCC CAGCAGGCTT CCAGCTCCTG 720  
GACCAGAGA TCTGTGGCGA CATTGATGAG TGCAAGGACC CAGATGCGTG CAGCCAGATG 780  
TGTGTCAAT ACAAGGGCTA TTTTAAGTGT GAGTGTACC CTGGTTCGGA GATGGACCTA 840  
CTGACCAAGA ACTGCAAGCG TCTGCTGGC AAGAGCCCAT CCCTAATCTT CACCAACCGC 900  
ACGAGTGGG AGGATCGACC TGTGAAGCGG AACTATTTCAC GCCTCATCCC CATGCTCAAG 960  
AATGTCTGG CACTAGATGT GGAAGTTGCC ACCAATCGCA TCTACTGGTG TGACCTCTCC 1020  
TACCGTAA GA TCTATAGCG CTACATGGAC AAGGCCAGTG ACCCGAAAGA GCGGAGGGT 1080  
CTCATTTAG AGCAGTTGCA CTCTCCAGAG GGCCTGGCAG TGGACTGGGT CCACAAGCAC 1140  
ATCTACTGA CTGACTCGGG CAATAAGACC ATCTCAGTGG CCACAGTTGA TGGTGGCGCG 1200  
CGAGCACTC TCTTCAGCG TAACCTCAGT GAACCCCGGG CCATCGCTGT TGACCCCTTG 1260  
CGAGGGTTCA TGTATTGGTC TGACTGGGGG GACCAGGCCA AGATTGAGAA ATCTGGGCTC 1320  
AACGGTGTG ACCGSCAAAC ACTGGTGTCA GACAATATTG AATGGCCCAA CGGAATCAC 1380  
CTGGATCTGC TGAGCCAGCG CTGTACTGCG GTAGACTCCA AGCTACACCA ACTGTCCAGC 1440  
ATTGACTTCA GTGAGGCAA CAGAAAGACG CTGATCTCCT CCAGTACTT CCGTGGCCAC 1500  
CCTTTTGGGA TAGCTGTGTT TGAGGACAAG GTGTTCTGGA CAGACCTGGA GAACGAGGCC 1560  
ATTTTCAGTG CAAATCGGCT CAATGGCTGT GAAATCTCCA TCCTGGCTGA GAACCTCAAC 1620  
AACCACATG TCAATGTGAT CTTCATGAG CTGAAGCAGC CAAGAGCTCC AGATGGCTGT 1680  
GAGCTGAGT TCCAGCTAA TGAGGGCTGT GAATACCTGT GCCTTCTGCT TCCTCAGATC 1740  
TCCAGCACT TCTATCTGCG CACATGTGCG TGTCTGACA CAATGTGGCT GGTCCAGAG 1800  
ATGAAGAGGT GCTACCGAGA TGCAATGAA GACAGTAAGA TGGGCTCAAC AGTCACTGCC 1860  
CGCTTTATG GGATCATCGT GCCCATAGTG GTGATAGCCC TCCTGTGCAT GAGTGGATAC 1920  
CTGATCTGGA GAAATCGAA CCGGAAGAAC ACCAAAGCA TGAATTTGA CAACCCAGTC 1980  
TACAGGAAA CAACAGAGA AGAAGATGAA GATGAGCTCC ATATAGGGAG AACTGTCTAG 2040  
ATTGGCATG TCTATCTGCG ACGAGTGGCA TTAAGCCTTG AAGATGATGG ACTACCTCTG 2100  
GGATGGGATC ACCCCCTTCG TGCCCTCATG AATTCAGTCC CATGCACTAC ACTCGGATG 2160  
GTGTATGACT GGATGAATGG GTTTCTATAT ATGGGTCTGT GTGAGTGTAT GTGTGTGTGT 2220  
GATTTTTTT TTTAAATTTA TGTTCGGGAA AGGTAACCAC AAAGTTATGA TGAAGTGC 2280  
ACATCCAAAG GATGTGAGAG TTTTCTATG TATAATGTT TATACACTTT TTAAGTGT 2340  
GCACTACCA CTGGAATTC GTGGAATGCC TACTGCTGAC TAACATGATG CACATAACCA 2400  
AATGGGGGCC AATGGCACAG TACCTTACTC ATCATTTAAA AACTATATTT ACAGAAGATG 2460  
TTGGTGTCT GGGGGGCTTT TTTAGGTTTT GGGCAITTTG TTTTGTAAA TAAGATGATT 2520  
ATGCTTTGTG GCTATCCATC AACATAAGT

Seq ID NO: 357 Protein sequence

Protein Accession #: NP\_059992

55  
60  
65  
70

1 11 21 31 41 51  
MGLPEPGPLR LLAALLLLLL LLLLRLQLHA AAAADPLLGG QGPAKECEKD QFQCRNERCI 60  
PSVWRCEDED DCLHSDDED CPKKTCAESD FTDNGHCII ERWKCDGEE CPDGSDESEA 120  
TCTKQVCPAE KLSCGPTSHR CVPASWRCDG EKDCGGADE AGCATSLGCT RGDEFQCGDG 180  
TCVLAIRKCN QEQDCPDGSD EAGCLQGLNE CLHNGGCSH ICTDLKIGFE CTCFAGFQLL 240  
DQKTCGDI DE CKDPDACSQI CVNYKGYFKC ECYPGCEMDL LTKNCKAAAG KSPSLIFTNR 300  
TSAEDRPVFR NYSRLIPMLK NVVALDVEVA TNRIYWCDSL YRKIYSAYMD KASDPKEREV 360  
LIDEQLHSPE GLAVDWVHKH IYWTDSGNKT ISVATVDGGR RRTLFSENLS EPRAIADVPL 420  
RGPMYNSDWG DQAKIEKSL NGVDRQTLVS DNIWPNIGIT LDLLSRLYN VDSKLHLQSS 480  
IDPSGGRNKT LISSTDPLSH PFGIAVFEDK VFWTDLENEA IFSANRLNGL EISILAENLN 540  
NPHDIVIFHE LKQPRAPDAC ELSVQPNNGC EYLCLPAPQI SSHSPKYTCA CPDTMWLGP 600  
MKRCYRDANE DSKMGSTVTA AVIGIIVPIV VIALLCMSGY LIWRNWKRRN TKSMNPDNPV 660  
YRKTTREDE DELHIGRTAQ IGHVYPARVA LSLEDDGLP

Seq ID NO: 358 DNA sequence

Nucleic Acid Accession #: M27826

Coding sequence: &lt;1-503

75  
80  
85

1 11 21 31 41 51  
AGCCCAAGAA ACATCTCACC AATTTCAAAT CTGATCTATT CGGCTTAGCG ACTGAAGATT 60  
GACGCTGCC GATCGCTCG GAAGTCCCTT GGACCATCAC AGAAGCCGAG CTTGGGTTAA 120  
CTCTCAGAGT GGAGGGTAAG TCCATCCCTT GTTTAATCGA TACGGGGGCT ACCCACTCCA 180  
CGTTGCTTCT TTTCAAGGG CCTGTTTCCC TTGCCCCAT AACTGTTGTG GGTATTGAGC 240  
GCCAAGCTTC AAAACCCCTG AAAACTCCCC CACTCTGGTG CCAACTTGA CAACACTCTT 300  
TTATGCACTC TTTTATGTT ATCCCCACCT GCCCACTTCC CTTATTAGGC CGAAATATTT 360  
TAACCAAAAT ATCTGCTTCC CTGACTATTC CTGGAGTACA GCTACATCTC ATTGCTGCCC 420  
TTCTTCCCAA TCCAAAGCCT CCTTTGTGTC CTCTAACATC CCCACAATAT CAGCCCTTAC 480  
CACAAAGACT CCTTTCAGT TAATCTCTCC CACTCTAGGT TCCCAAGCGG CCCCTAATCC 540  
CACTTGAAGC AGCCCTGAGA AACATGCCCC ATTCTCTCTC CATACCAACC CCCAAAATTT 600  
TTGCGCGCTC CAACACTTCA ACACTATTTT GTTTTATTTG TCTTATTAAT ATCAGAAGGC 660



AGGAATGTCA GGCCTCTGAG CCCAGGCCAG GCCATCGCAT CCCCTGTGAC TTGCACGTAT 720  
 ACATCCAGAT GGCCTGAAGT AACTGAAGAT CCACAAAGA AGTAAAAACA GCCTTAACGT 780  
 ATGACATTCC ACCATTGTGA TTTGTTCTGT CCCACCCCTA ACTGATCAAT GTACTTTGTA 840  
 ATCTCCCCCA CCTTAAAGAA GGTTCCTTGT AATTCTCCCC ACCCTTGAGA ATGACTTTTG 900  
 TGAGATCCAC CCCTGCCAC CAGAGAACAA CCCCTTTGA TTGTAATTTT TTATTACCTT 960  
 CCCAAATCCT ATAAACAGC CCCACCCCTA TCTTCTTCA CTGACTCTCT TTTCGGACTC 1020  
 AGCCACCGGC ACCCAGGTGA AATAAACAGC TTTATTGCTC AC

Seq ID NO: 359 Protein sequence  
 Protein Accession #: AAA65999

1 11 21 31 41 51  
 PKKHLTNPKS DLFLGLATEDW RCPIASEVPM TITEAELRVT LTVEGKSIPC LIDTGATHST 60  
 LPSFQGFVSL APITVVIGID QASKPLKTPP LWCQLGQHSF MHSFLVIPTC PLFLLGRNLL 120  
 TKLSASLTIP QVQLHLIALA LPNFKPPLCP LTSFQVQPLP QDLPSA

Seq ID NO: 360 DNA sequence  
 Nucleic Acid Accession #: NM\_001854  
 Coding sequence: 162-5582

1 11 21 31 41 51  
 AACCATCAAA TTTAGAAGAA AAAGCCCTTT GACTTTTTCC CCCTCTCCCT CCCCAATGGC 60  
 TGTGTAGCAA ACATCCCTGG CGATACCTTG GAAAGGACGA AGTTGGTCTG CAGTCGCAAT 120  
 TTGTTGGGTT GAGTTCACAG TTGTGAGTGC GGGGCTCGGA GATGGAGCCG TGGTCCCTTA 180  
 GGTGGAAAAC GAAACGGTGG CTCTGGGATT TCACCGTAAC AACCTTCGCA TTGACCTTCC 240  
 TCTTCCAAGC TAGAGAGGTG AGAGGAGCTG CTCAGTTGA TGTACTAAAA GCACCTAGATT 300  
 TTCACAATTC TCCAGAGGGA ATATCAAAAA CAACGGGATT TTGCACAAAC AGAAAGAAAT 360  
 CTAAGGGCTC AGATACTGCT TACAGAGTTT CAAAGCAAGC ACAACTCAGT GCCCAACAAA 420  
 AACAGTTATT TCCAGTGTGA ACTTCCAG AAGACTTTTC AATACTATT ACAGTAAAC 480  
 CAAAAAAGG AATTCACTCT TCCCTTTTAT CTATATATAA TGAGCATGGT ATTCAGCAA 540  
 TTGGTGTGTA GGTGGGAGA TCACCTGTTT TTCTGTTTGA AGACCACT GGAAAACCTG 600  
 CCCAGAGA CTATCCCTTC TCCAGAACTG TTAACATCGC TGACGGGAAG TGGCATCGGG 660  
 TAGCAATCAG CGTGAGAAAG AAAACTGTGA CAATGATTGT TGATTGTAAG AAGAAAACCA 720  
 CGAAACCACT TGATAGAAGT GAGAGAGCAA TTGTTGATAC CAATGGAAAT ACGGTTTTG 780  
 GAACAGGAT TTTGATGAA GAAGTTTTG AGGGGACAT TCAGCAGTTT TTGATCAGAG 840  
 GTGATCCCAA GGCAGCATAT GACTACTGTG AGCATATAG TCCAGACTGT GACTCTTCAG 900  
 CACCAAGGC TGCTCAAGCT CAGGAACCTC AGATAGATGA GTATGCACCA GAGGATATAA 960  
 TCGAATATGA CTATGAGTAT GGGGAAGCAG AGTATAAAGA GGCTGAAAGT GTAACAGAGG 1020  
 GACCCACTGT AACTGAGGAG ACAATAGCAC AGACGGAGGC AAACATCGTT GATGATTTTC 1080  
 AAGAATACAA CTATGGAGCA ATGGAAGATT ACCAGACAGA AGCTCCTAGG CATGTTTCTG 1140  
 GGACAAATGA GGCAGATCCA GTTGAAGAAA TATTTACTGA AGAATATCTA ACGGAGAGG 1200  
 ATTATGATT CAGAGGAAA AATTCTGAGG ATACACTATA TGAACAACAA GAAATAGACG 1260  
 CAGGGATTTC TGATCTTCTG GTAGATGGAG ATTTAGGCGA ATATGATTTT TATGAATATA 1320  
 AAGAATATGA AGATAAACCA ACAAGCCCC CTAATGAAGA ATTTGGTCCA GGTGTACCAG 1380  
 CAGAACTGA TATTACAGAA ACAAGCATAA ATGGCCATGG TGCAATGGA GAGAAAGGAC 1440  
 AGAAAGGAGA ACCAGAGTGC GTTGAAGCTG GTATGCTTGT CGAAGGACCA CCAGGAGCAG 1500  
 CAGGACCTGC AGGTATTATG GGTCTCTCAG GTCTACAAAG CCCCACTGGA CCCCTCGGTG 1560  
 ACCCTGGCGA TAGGGGCCCC CCAGGACGTC CTGGCTTACC AGGGGCTGAT GGTCTACCTG 1620  
 GTCTCTCTGG TACTATGTTG ATGTTACCGT TCCGTTATGG TGGTGATGGT TCCAAAGGAC 1680  
 CAACCATCTC TGCTCAGGAA GCTCAGGCTC AAGCTATTCT TCAGCAGGCT CGGATTGCTC 1740  
 TGAGAGGCCA CCTGGGCCA ATGGGTCTAA CTGGAAGACC AGGTCTGTG GGGGGGCGTG 1800  
 GTTCACTCGG GGCACAAAGT GAGATGGTGT ATCCAGGTCC TCAGGGCCCT CGAGGCGTCC 1860  
 AGGGTCCCCC TGGTCCAACG GGAAAACCTG GAAAAGGGGG TCGTCCAGT GCAGATGGAG 1920  
 GAAGAGGAAT GCCAGGAGAA CCTGGGCGAA AGGGAGATCG AGGGTTGAT GGACTTCCGG 1980  
 GTCTGCCAGG TGCAAAAGST CACAGGGGTG AACGAGGTCC TCAAGGTCTT CCAGGTCTCT 2040  
 CTGGTGATGA TGGAAATAGG GGAGAGATG GAGAAATTGG ACCAAGAGGT CTTCCAGGTG 2100  
 AAGCTGGCCC ACGAGTTTGT CTGGGTCCAA GGGGAATCC AGGAGCTCCA GGGCAGCCTG 2160  
 GTATGGCAGG GTGATAGGCT CCCCAGGAC CAAAAGGGAA CATGGTCCC CAAGGGAGC 2220  
 CTGGGCTCTC AGGTCAACAA GGAATCCAG GACCTCAGGG TCTTCTGTGT CCACAAGGTC 2280  
 CAATTGGTCC TCTGTGTGAA AAAGGACCAC AAGGAAAACC AGGACTTGCT GGACTTCTCT 2340  
 GTGCTGATGG GCCTCTGCTG CATCTGGGA AAGAAGGCCA GTCTGGAGAA AAGGGGGCTC 2400  
 TGGGTCCCCC TGGTCCACAA GGTCTATTG GATNNCCGG CCCCCGGGGA GTAAAGGGAG 2460  
 CAGATGGTGT CAGAGGTCTC AAGGGATCTA AAGGTGAAA GGGTGAAGAT GGTTTTCCAG 2520  
 GATTCAAAGG TGACATGGGT CTAAAAGGTG ACAGAGGAGA AGTTGGTCAA ATTGGCCCAA 2580  
 GAGGNAAGA TGGCCTGAA GGAACCAAG GTCCAGCAGG CCAACTGGA GACCCAGGTC 2640  
 CTTCAAGTCA AGCAGGAGAA AAGGAAAAC TTGGAGTTCC AGGATTACCA GGAATCCAG 2700  
 GAAGACAAGG TCCAAAGGAT TCCATGGAT TCCCTGGGTT TCCAGGTGCC AATGGAGAGA 2760  
 AAGGTGCAAG GGGAGTAGCT GGCAAAACAG GCCCTCGGG TCAGCGTGGT CCAACGGGTC 2820  
 CTCGAGGTTC AAGAGGTGCA AGAGGTCCCA CTGGGAAAAC TGGGCCAAAG GGCACTTCAG 2880  
 GTGGCGATGG CCTCTCTGCC CCTCCAGGTG AAAGAGGTCC TCAAGGACCT CAGGGTCCAG 2940  
 TTGATTCCCT TGGACCAAAA GGCCCTCTCT GACCACAGG AAGGATGGGC TGCCCGAGAC 3000  
 ACCCTGGGCA ACGTGGGAG ACTGGATTTC AAGSCAAGAC CGGCCCTCTT GGGCCAGGG 3060  
 GAGTGGTTGG ACCACAGGGA CCAACGGGTG AGACTGGTCC AATAGGGGAA CGTGGGTATC 3120  
 CTGGTCTCTC TGGCCTCTCT GGTGAGCAAG GTCTTCTG TGCTGCAGGA AAAGAAGGTG 3180  
 CAAAGGTGTA TCCAGGTCTC CAAGGTATCT CAGGGAAGA TGGACCAAGA GGAATTACGTG 3240  
 GTTTCACAGG GGAAGAGGT CTCTCTGGAG CTCAGGTGTC ACCTGGAAGT AAAGAGGGG 3300  
 AAGGTCCCA GGGCCACCA GGTCCAGTTG GCTCACCAG AGAACGTGGG TCAGCAGGTA 3360  
 CAGCTGCCCC AATTGGTTTA CGAGGGGCC CGGACCTCA GGTCTCTCT GGTCCAGCTG 3420  
 GAGAGAAAGG TGCTCTGGA GAAAAGGTG CCAAGGGGCC TGACAGGAGA GATGAGGTT 3480  
 AAGGTCTCTG TGGTCTGGA GGGCCAGCTG GTCTCTGGG CTCCCTGGG GAAGACGGAG 3540  
 ACAAGGTGTA AATTGGTGA CCGGACAAA AAGGCAGCAA GGGTGGCAAG GGAGAAAATG 3600  
 GCCCTCCCG TCCCTCAGG CTTCAGGAC CAGTTGGTGC CCTGGAATT GCTGAGGTTG 3660  
 ATGGTGAACC AGSTCTAGA GAACAGCAGG GGATGTTGG GCAAAAAGGT GATGAGGGTG 3720  
 CCAGAGGCTT CCTGGACCT CTTGGTCCAA TAGGTCTTCA GGGTCTGCA GGGCCACCTG 3780  
 GTGAAAAGG TGAATGGG GATGTTGGT CATGGGGGCC ACCTGGTCTT CCAGGCCCAA 3840



5  
10  
15  
20  
25  
30  
35  
40

```

GAGGCCCTCA AGGTCCCAAT GGAGCTGATG GACCACAAGG ACCCCAGAGT TCTGTTGGTT 3900
CAGTTGGTGG TGTGGAGAG AAGGGTGAAC CTGGAGAAGC AGGAAACCCA GGGCCTCCTG 3960
GGGAAGCAGG TGTAGGCGGT CCCAAAGGAG AAAGAGGAGA GAAAGGGGAA GCTGGTCCAC 4020
CTGGAGCTGC TGGACTCTCA GGTGCCAAGG GGCCGCCAGG TGATGATGGC CTAAGGGTGA 4080
ACCCGGGTCC TGTGGTGTTC CTGGAGATC CTGGTCTCC TGGGGAACTT GGCCTCGAG 4140
CTCAAGATGG TGTGGTGGT GACAAGGGTG AAGATGGAGA TCTGGTCAA CCGGTCCTC 4200
CTGGCCATC TGGTGAGGCT GGCCACCAG GTCTCTCTGG AAAACGAGGT CCTCTGGAG 4260
CTGACAGTGC AGAGGGAGAA CAAGGTGAAA AAGGTGCTAA GGGGAAGCA GGTGCAGAA 4320
GTCTCTCTGG AAAAACCGGC CCAAGTGGTC CTCAGGAGC TGCAGGAAAG CTTGGTCCAG 4380
AAGGTCTTGG GGCATCCCT GGTCTGTGG GAGAACAAGG TCTCCTGGA GCTGCAGGCC 4440
AAGATGGACC ACCTGGTCTT ATGGGACCTC CTGGCTTACC TGGTCTCAA GGTGACCTG 4500
GCTCCAAGGG TGAAGAGGA CATCTGGTT TAATTGGCTT GATTGGTCT CAGAGAGAAC 4560
AAGGGGAAAA AGGTGACGGA GGGCTCCCTG GAATCAAGG ATCTCCAGG GCAAAGGGG 4620
ATGGGGGAAAT TCTGGTCTCT GCTGGTCTCT TAGGTCCACC TGGTCTCCA GGCTTACCAG 4680
GTCTCTCAAG CCCAAAGGGT AACAAAGGCT CTACTGGACC CGCTGGCCAG AAAGGTGACA 4740
GTGGTCTTCC AGGGCTCTCT GGGCTCCAG GTCCACCTGG TGAAGTCATT CAGCCTTTAC 4800
CAATCTTGTCT CTCCAAAAAA ACGAGAAGAC ATACTGAAG CATGCAAGCA GATGCAGATG 4860
ATAATATTCT TGATTACTCG GATGGAATGG AAGAAATATT TGGTTCCTC AATTCCCTGA 4920
AACAGACAT CGAGCATATG AAATTTCCAA TGGGTACTCA GACCAATCCA GCCCGAACTT 4980
GTAAAGACCT GCAACTCAGC CATCTGACT TCCAGATGG TGAATATTGG ATTGATCCTA 5040
ACCAAGGTTG TCCAGGAGAT TCCTTCAAAG TTTACTGTAA TTTACATCT GGTGGTGAGA 5100
CTTGCAATTA TCCAGACAAA AAATCTGAGG GAGTAAGAA TTCTCATGG CCAAGGAGA 5160
AACCAGGAAG TTGGTTTAGT GAATTTAAGA GGGGAAAACT GCTTTCATAC TAGATGTTG 5220
AAGGAAATTC CATCAATATG GTGCAATGA CATCTCTGAA ACTTCTGACT GCCTCTGCTC 5280
GGCAAAATTT CACCTACACC TGTCATCAGT CAGCAGCCTG GTATGATGTG TCATCAGGAA 5340
GTTATGACAA AGGCCTCTCG TTCTGGGAT CAAATGATGA GGAGATGTCC TATGACAATA 5400
ATCCTTTTAT CAAACACTGT TATGATGGT GTAAGTCCAG AAAAGGCTAT GAAAAAACTG 5460
TCATTGAAAT CAATACACCA AAAATTGATC AAGTACCTAT TGTGATGTC ATGATCAGT 5520
ACTTTGGTGA TCAGAAATCAG AAGTTCGGAT TTGAAGTGG TCTGTTTGT TTTCTGGCT 5580
AAGATTAAAG CAAAGAACAT ATCAAATCAA CAGAAAAATG ACCTTGGTGC CACCAACCCA 5640
TTTTGTGCCA CTAAGTGTG TTGAATAAGG ATGATGGAA AACACGCTG CATATACAGG 5700
TACCATTTAG GAAATACCGA TGCCTTGTG GGGGAGAAAT CACAGACAAA AGCTTTGAAA 5760
ATCATAAGA TATAAGTTGG TGTGGCTAAG ATGGAACAG GGCTGATTCT GTATTCCAA 5820
TTCTCAACTC TCCTTTTCTT ATTTGAAATT CTTTGGTCT GTAGAAAACA AAAAAGAAA 5880
AATATATATT CATAAAAAAT ATGGTGCTCA TTCTCATCCA TCCAGGATG ACTAAAACAG 5940
TGTGTTTAA TAAATGTAT TATTTGTGT ACAGTTCTAT ACTGTTATCT GTGTCCATT 6000
CCAAACTTG CAGCTGTCCC TGAATTCGC TGACTCTAAT TTATGAGGAT GCGCAACTCT 6060
GATGGCAATA ATATATGTAT TATGAAATG AAGTTATGAT TTCCGATGAC CCTAAGTCCC 6120
TTCTTTGGT TAATGATGAA ATTCCTTTGT GTGTGTTT

```

Seq ID NO: 361 Protein sequence  
Protein Accession #: NP\_001845

45  
50  
55  
60  
65  
70  
75

```

1 11 21 31 41 51
MEPWSRWNK KRWLWDFVT TLALTFLEQA REVGAAPVD VLKALDFENS PEGISKTTGF 60
CTNRKNSKGS DTAYRVSKQA QLSAPTQQLF PGSTFPEDFS ILFTVKPKKG IQSPLLSIYN 120
EHGIQOIGVE VGRSPVPFLE DHTGKEPAPD YPLFRTVNIA DGKHWRAIS VEKKTIVMIV 180
DCKKTKTKPL DRSERAIVDT NGITVFGTRI LDEEVFEGDI QQFLITGDPK AAYDYCEHYS 240
PDCDSSAPKA AQAEQEPQIDE YAPEDIIEYD YEYGEAEYKE ABSVTEGTV TESTIAQTEA 300
NIVDDFQSYN YGTMESYQTE APRHVSQTN PNPVEEIFTE EYLTGEDYDS QRKNSDITLY 360
ENKEIDGRDS DLLVDGLDGE YDFYKEYEY DKPTSPFNNE FPGVPASTD ITETSINGHG 420
AYGEKQKGE PAVVEFGMLV EGPPGAPGA GIMGPPGLQG PTGPPGDPGD RGPPGRPLP 480
GADGLPGPPG TMLMLPFYRG GDGSKGPTIS AQEAQAAIL QQAIALRGP PGPMGLTRP 540
PVGPGPGSSG AKGESQDPGP QGPRGVQGP GPTGKPKRG RPADGGRGM PGPBGAKGDR 600
GPDGLPLPLP DKHHRGERG QGPPGPPGDD GMRGEDGEIG PRGLPGEAGP RGLLPGRGT 660
GAPGQPGNAG VDGPPGPKGN MGPPQGEPPG GQQGNPGPQG LPGPQQPIGP PGEKGPQKGP 720
GLAGLPGADG PGPHPKKGSG SGEKALGPP GPQQPIGXPG PRGVKADGV RGLKSGKGEK 780
GEDGPFPGFK DMGLKGRDGE VQIGFRGXD GPEGPKGRAG PTGDPGPGSQ AGEKGLGV 840
GLPGYPGRQG PKGSTGPFPG FGANGKGRG GVAGKPGPRG QRGFTGPRGS RGARGPTGKP 900
GPKGTSGDGD PPGPPGERGP QGPPGPPGDD GMRGEDGEIG PRGLPGEAGP RGLLPGRGT 960
GPPGPGGVVG PQGTGTGTGP IGERGYPGPP GPPGQGLPG AAGKEGAKGD PGPGISGKD 1020
GPAGLRGPPG ERGLPGAQGA PGLKGGEGPQ GPPGVPVSPG ERGSAGTAGP IGLRCRPGPQ 1080
GPPGPAGEKG APGEKGPQGP AGRDGVQGPV GLPGPAGPAG SPGEDGDKGE IGEPGQKGSK 1140
GGKGENPPPG PPQLQGPVGA PGIACGDGEP GPRGQQGMFG QKDEGARGF PGPPGPIGLQ 1200
GLPGPPGEKG ENGDIVPMGP PGPGRGPRGP GPNGADGPPG PPGSVGSVGG VGEKGEPEGA 1260
GNPPGPEAG VGGPKGERGE KGEAGPPGAA GPPGAKGPPG DDGPKGNPVP VGGPDGPPG 1320
GELGPAQDGD VGGDKGEDGD PQQPPGPPGS GEAGPPGPPG KRGPPGAAGA EGRQGEKGA 1380
GEAGAEPPPG KTGVPVQGP AGKPGPEGLR GIPGVPVGEQG LPGAAGDGP PGPMGPPGLP 1440
GLKGDPSKGS EKGHPGLIGL IGPPGEQGEK GDRGLFGTQG SPGAKDGDI PGPAQPLGPP 1500
GPPGLPFPQG PKGNKSTGTP AGQKGDGSLP GPPGPPGPPG EVIQPLPILS SKKTRRTEG 1560
MQADADDNII DYSDGMEEIF GSLNLSKQDI EHKFPFMGTQ TNPARTCKDL QLSHPDFPDG 1620
EYWDPNQGC SGDSFKVYCN PFSGETCTY PDKKSEGVRI SSWPKKPGS WFSFKRGL 1680
LSYLDVEGNS INVMQTFLEK LITASARQNF TYEHCQSAW YDVSSGSYDK ALRPLGSNDE 1740
EMSYDNNPFI KTYLDGCTSR KYEKTIVIEI NTFKIDQVPI VDMISDFGD QNKQFGEV 1800
PVCFLG

```

Seq ID NO: 362 DNA sequence  
Nucleic Acid Accession #: NM\_003107  
Coding sequence: 351-1775

85

```

1 11 21 31 41 51
TTCCCCAGCA TTGAGAAAC TCCTCTCTAC TTTAGCACGG TCTCCAGACT CAGCCGAGAG 60
ACAGCAAACT GCAGCGCGGT GAGAGAGCGA GAGAGAGGGA GAGAGAGACT CTCCAGCCTG 120
GGAATATAA CTCCTCTGCG AGAGGCGGAG AACTCCTTCC CCAATCTTT TGGGACATTT 180

```

TCTCTCTTTA CCCACCTCG CCCCTGCGAG GAGTTGAGGG GCCAGTTGCG CCGCCGCGCG 240  
 CGTCTTCCCG TTCCGGGTGTG GCTTGGCCCG GGGAAACCGG AGGGCCCGCG GATCGCGCGG 300  
 CGCGCGCGCG GAGGGTGTGA GCGCGCGTGG GCGCCCGCGG AGCGCGAGGC ATGGTGCAGC 360  
 AAACCAACAA TGCCGAGAAC ACGGAAGCGC TGCTGGCCCG CGAGAGCTCG GACTCGGGCG 420  
 CGCGCTCGA GCTGGGAATC GCCTCTCCCG CCACGCCCGG CTCACCGGCC TCCACGGCGG 480  
 GCAAGGCGGA CGACCCGAGC TGGTGCAAGA CCCCGAGTGG GCACATCAAG CGACCCATGA 540  
 ACGCCTTCAT GGTGTGTGCG CAGATCGAGC GCGCGAAGAT CATGGAGCAG TCGCCCGACA 600  
 TGCAACAACG CGAGATCTCC AAGCGGCTGG GCAAAACGCTG GAAGCTGCTC AAAGACAGCG 660  
 ACAAGATCAA TTTCATTGGA GAGGCGGAGC GGCTGGCCCT CAAGCACATG GCTGACTACC 720  
 CGGACTACAA GTACCGGCCG AGGAAGAAGG TGAAGTCCGG CAACGCCAAC TCCAGCTCCT 780  
 CGCGCGCGCG CTCTCTCAAG CCGGGGAGGA AGGAGACAA GGTCCGTGGC AGTGGCGGGG 840  
 GCGGCCATGG GGGCGCGCGC GCGCGCGGGA GCAGCAACGC GGGGGAGGGA GCGCGCGGTG 900  
 CGAGTGGCGG CGCGCGCAAC TCCAAACCGG CGCAGAAAAA GAGCTCGGCG TCCAAAGTGG 960  
 CGGGCGCGCG GGTCTGTGGG GTTAGCAAAAC CGCACGCCAA GCTCATCTCG GCAGGCGGCG 1020  
 GCGCGCGCGG GAAAGCAGCG GCTCGCCCGG CCGCTCTCTT CGCGCGGAA CAGGCGGGGG 1080  
 CGCGCGCGCT GCTGCGCTCG GCGCGCGCGG CGACCAACCA CTGCTGTAC AAGGCGCGGA 1140  
 CTCCAGCGCG CTGCGCTCTC GCCTCTCTCG CAGCTCTCGG CTCGCGCAGG CTCGCGGGCC 1200  
 CGGGCAAGCA CCTGGCGGAG AAGAAGGTGA AGCGGTCTA CTGTTCGGCG GCGCTGGGCA 1260  
 CGTGTGTGTC GCGCGTGGCG GCGGTGGCGG CGGAGCGCGA CCGCAGCGAG CCGCTGGGCG 1320  
 TGTAAGAGGA GAGGGGCGCG GGCTGTCTCG CCGACGCGCG CAGCTGAGC GCGCGCAGCA 1380  
 GCGCGCGCTC GTCCCGCGCG GCGCGCGCGT CGCGCGCGA CCACCGCGCG TACGCGAGCC 1440  
 TGCGCGCGCG CTGCGCGCGC CCGTCCAGCG CGCGCTCGCA CGGCTCTCTC TCGGCTCTGT 1500  
 CCCACTCTCT CTCTCTCTCT TCCTCGGGCT CCTCGTCTCT CGACGACGAG TTCGAAGACG 1560  
 ACCTGTCTGA CCTGAACCCC AGCTCAAACT TTGAGAGCAT GTCCCTGGGC AGCTTCAGTT 1620  
 CGTGTGTGCG GCTCGACCGG GACTCGGATT TTAACCTCGA GCGCGGCTCC GGCTCGCACT 1680  
 TGAGTTTCCC GGACTCTCTG ACGCCGAGG TGAGCGAGAT GATCTCGGGA GACTGGCTCG 1740  
 AGTCAGCAT CTCAACCTG GTTTTCACT ACTGAAGGGC GCGCAGGCG GAGAGAGGGC 1800  
 CGGGGGGGGT AGGAGAGGAG AAAAAAAGG TGAATAAAG AAACGAAAG GACAGACGAA 1860  
 GAGTTTAAAG AGAAAAAGGA AAAAAGAAAG AAAAAGTAAG CAGGGCTCGT TCGCCCGCGT 1920  
 TCTGTGTGTC GGATCAAGGA GCGCGCGCGC GTTTTGGACG CGCGCTCCCA TCCCGCCACT 1980  
 TCCCGCGCGC GGGACCCACT CTGCGCAGCG GGAGGAGCGC GAGGAGGAA GAGGGTAGAC 2040  
 AGGGGCGAAC TGTGATTGTT GTTATTGATG TTGTTGTTGA TGGCAAAAAA AAAAAGCGAC 2100  
 TCGAGTTTGT CTCCCTTTTG CTGGAAGAGA CCGCTCTCCC CTTCACGGA GCTTCCGAC 2160  
 TTGTCTGCAC CCGCAGCAAG AAGCGAGTT AGTTTTCTAG AGACTTGAAG GAGTCTCCCC 2220  
 CTCTCTGCAT CACCACTCTG GTTTTGTGTT ATTTTGTCTC TTGTTCAAGA AAGGAGGAGA 2280  
 GAAACCGAGG GAACCCCTCC CCGCTTTTTC TAAACGCGTG ATGAAGACAG AAGGCTCCGG 2340  
 GGTGACGAAT TTGGCGGATG CGAGATGTTT TGGGGGAGCG CCGGAGCTGA GAGACTCCAC 2400  
 GCAGGCGAAT TCCGTTTGG GCGCTTTTTC TCCTCCCTCT TTTCCCTTGT CCGCTCTGCG 2460  
 AGCGGAGAGA GGAGATGTTG AGGGAGGAG GCCAGCCAGT GTGACCGCGG CTAGGAAATG 2520  
 ACCCGAGAAC CCGGTTGGAA GCGCAGCAGC GGGAGCTAGG GCGCGGGGCG GAGGAGGACA 2580  
 GAACTGGAA GGGGGTTTAC GGTCAAACTG AATGGATTG GCACGTTGGG GAGCTGCGCG 2640  
 CGGCGCGTGC TGGGCTCCG CTTTCTTTTC TACGTGAAAT CAGTGAGGTG AGACTTCCCA 2700  
 GACCCGAGG GCGTGAGGA GAGGAGACTG TTTGATGTGG TACAGGGGCA GTCAGTGGAG 2760  
 GCGAGTGTGT TTCGAAAAA AAAAAGAAA AAAAGGG

Seq ID NO: 363 Protein sequence  
 Protein Accession #: NP\_003098

1 11 21 31 41 51  
 MVQQTNNAEV TEALLAGESS DSGAGLELGI ASSPTPGSTA STGGKADDPs WCKTPSGHIK 60  
 RPNNAFVWVS QIERRKIMQD SPDMHNAEIS KRLGKRWKLL KDSDKIPPIR EAERLRLKHM 120  
 ADYDPYKYRP RKVYKSGNAN SSSSAAASSK PGEKGDVKGV SGGGGHGGGG GGGSSNAGGG 180  
 GGGAGGGGAN SKPAQKSCG SKVAGGAGGG VSKPHAKLIL AGGGGGGKAA AAAAASPAAE 240  
 QAGAAALLPL GAAADHSLY KARTPSASAS ASSAASASAA LAAPGKHLAE KVKRVYVLF 300  
 GLGTSSSPVG GVGAGADPSD PLGLYEELGA GCSFDAPSLs GRSSAASSPA AGRSPADHRG 360  
 YASLRAASPA PSSAPSHASS SASSHSSSSS SSSSSSSDDE FEDDLLDLNP SSNFESMSLG 420  
 SPSSSALDR DLDFNFEFGS GSHFEFPDYC TPEVSEMISS DWLESISNL VFTY

Seq ID NO: 364 DNA sequence  
 Nucleic Acid Accession #: U10860  
 Coding sequence: 123-2204

1 11 21 31 41 51  
 TGCGCGCTGC TCCTGACCA GCGCTCTCTC TCAACCTCAG CCGCGCGCGC CGACCCCTTCC 60  
 GGCACCTCTC CGCCCGCTCT GCTACTGTGC CCGTCAACGC CGCGGCTCCG GCCCTGGCCC 120  
 CGATGGCTCT GTGCAACGGA GACTCCAAGC TGGAGAATGC TGGAGGAGAC CTTAAGGATG 180  
 GCCACCACCA CTATGAAGGA GCTGTTGTCA TTCTGGATGC TGGTGCTCAG TACGGGAAAG 240  
 TCATAGACCG AAGAGTGAGG GAACTGTTCG TGCACTCTGA AATTTTCCCC TTGGAACAC 300  
 CAGCATTGTC TATAAAGGAA CAAGGATTCC GTGCTATTAT CATCTCTGGA GGACCTAATT 360  
 CTGTGTATGC TGAAGATGCT CCTGGTTTG ATCCAGCAAT ATTCACTATT GGCAGCCTG 420  
 TTCTTGAAT TTGCTATGTT ATGCAGATGA TGAATAAGGT ATTTGGAGGT ACTGTGCACA 480  
 AAAAAAGTGT CAGAGAAGAT GGAGTTTTC ACATTAGTGT GGATAATACA TGTTCAATTAT 540  
 TCAGGGGCGCT TCAGAAGGAA GAAGTTGTTT TGCTTACACA TGGAGATAGT GTAGACAAAG 600  
 TAGCTGATGG ATTCAGGTT GTGGCAGGTT CTGGAACAT AGTAGCAGGC ATAGCAAAATG 660  
 AATCTAAAAA GTTATATGGA GCACAGTTC ACCCTGAAGT TGGCCTTACA GAAATGGA 720  
 AAGTAATACT GAAGAATTTC CTTTATGATA TAGCTGATG CAGTGGAAAC TTCACCGTGC 780  
 AGAACAGAGA ACTTGAGTGT ATTCAGAGGA TCAAGAGAG AGTAGGCAGC TCAAAAGTTT 840  
 TGGTTTTACT CAGTGGTGGG GTAGACTCAA CAGTTTGTAC AGCTTTGTCT AATCGTGCTT 900  
 TGAACCAAGA ACAAGTCATT GCTGTGCACA TTGATAATGG CTTTATGAGA AAACGAGAAA 960  
 GCCAGTCTGT TGAAGAGGCC CTCAAAAGC TTGGAATTC GGTCAAAGT ATAAATGCTG 1020  
 CTCAATCTTT CTACATGGA ACAACAACC TACCAATATC AGATGAAGAT AGAACCCAC 1080  
 GGAAGAGAA TAGCAAAACG TTAATATGA CCACAAGTCC TGAAGAGAAA AGAAAAATCA 1140  
 TTGGGGATAC TTTTGTAAAG ATTGCCAATG AAGTAATTGG AGAATGAAC TTGAACCCAG 1200  
 AGGAGGTTTT CCTTGCCCAA GGTACTTTAC GGCTGTATCT AATTGAAAGT GCATCCCTTG 1260

TTGCAAGTGG CAAAGCTGAA CTCATCAAAA CCCATCAAAA TGACACAGAG CTCATCAGAA 1320  
 AGTTGAGAGA GGAGGGAAAA GTAATAGAAC CTCTGAAAAG TTTTCATAAA GATGAAGTGA 1380  
 GAAATTTGGG CAGAGAACTT GGACTTCCAG AAGAGTAGT TTCCAGGCAT CCATTTCCAG 1440  
 5 GTCCTGGCCT GGCAATCAGA GTAATATGTG CTGAAGAACC TTATATTGTG AAGGACTTTC 1500  
 CTGAAACCAA CAATATTTTG AAAATAGTAG CTGATTTTTC TGCAAGTGT AAAAAGCCAC 1560  
 ATACCCTATT ACAGAGAGTC AAAGCCTGCA CAACAGAAGA GGATCAGGAG AAGCTGATGC 1620  
 AAATTACCAG TCTGCATTCA CTGAATGCCT TCTTCTGCCC AATTAAACT GTAGGTGTGC 1680  
 AGGGTGACTG TGTTCCTTAC AGTTACGTGT GTGGAAATCTC CAGTAAAGAT GAACCTGACT 1740  
 10 GGAATCACT TATTTTCTG GCTAGGCTTA TACCTCGCAT GTGTCAACAC GTTAACAGAG 1800  
 TTGTTTATAT ATTGGCCCA CAGTTAAAG AACCTCCTAC AGATGTTACT CCCACTTCT 1860  
 TGACAACAGG GGTGCTCAGT ACTTTACGCC AAGCTGATT TGAGGCCCAT AACATTCTCA 1920  
 GGGAGTCTGG GTATGCTGG AAAATCAGCC AGATGCCGT GATTTTGACA CCATTACATT 1980  
 TTGATCGGGA CCCACTTCAA AAGCAGCCTT CATGCCAGAG ATCTGTGGTT ATTGGAACCT 2040  
 15 TTATTACTAG TGACTTCACT ACTGGTATAC CTGCAACACC TGGCAATGAG ATCCCTGTAG 2100  
 AGGTGATATT AAGATGGTC ACTGAGATTA AGAAGATTCC TGGTATTCT GGAATTATGT 2160  
 ATGACTTAAC ATCAAGCCC CCAGGAATA CTGAGTGGGA GTAATAAAT TC

Seq ID NO: 365 Protein sequence

Protein Accession #: AAA60331

1 11 21 31 41 51  
 MALCNGDSLK ENAGGDLKDG HHHYEGAVVI LDAGAQQYKV IDRRVRELTV QSEIFPLETP 60  
 25 AFAIKEQGFRI AIIISGGFNS VYAEDAPWFD PAIFITIGKPV LGICYGMQMM NKVFGGTVHK 120  
 KSVREDGVFN ISVDNTCSLF RGLQKEEVVL LTHGDSVDKV ADGFKVVARIS GNTVAGIANE 180  
 SKKLVGAGFH PEVGTCTGK VILKNFLYDI AGCSGTFTVQ NRELECIRES KERVGTSLK 240  
 VLLSGVDST VCTALLNRAL NQEQVIAVHI DNGFMRKRES QSVREALKKL GIOVKVINA 300  
 HSFYNGTTTL FISDEDRTPR KRISKTLNMT TSPEKRII GDTFVKIANE VIGEMNLKPE 360  
 30 EVFLAQGTLR PDLIESASLV ASGKAELIKT HNDTELIRK LREEGKVIIE LKDFHKDEVR 420  
 ILLRELGLPE ELVSRHPFPF PGLAIRVICA BEPYICRDFP ETNNILKIVA DFSASVKKPH 480  
 TLLQKVAKT TEEDQEKIMF ITSLSLNAF LLPIKTVGVQ GDCRSYSYVC GISSKDEPDW 540  
 ESLIFLARLI PRMCHNVRV VYIFGPPVKE PPTDVTPTFL TTGVLSTLRQ ADPEARNILR 600  
 ESGYAGKISQ MPVILTPLHF DRDPLQKQPS QRSVVIRTP ITSDFMTGIP ATPGNEIPVE 660  
 35 VVLKMTVEIK KIPGISRIMY DLTSKPPGTT EWE

Seq ID NO: 366 DNA sequence

Nucleic Acid Accession #: NM\_004219

Coding sequence: 46-654

1 11 21 31 41 51  
 GGGCCCTCAG ATGAATGCGG CTGTTAAGAC CTGCAATAAT CCAGAATGGC TACTCTGATC 60  
 45 TATGTTGATA AGGAAAATGG AGAACCAAGC ACCCGTGTGG TTGCTAAGGA TGGGCTGAAG 120  
 CTGGGGCTCG GACCTTCAAT CAAAGCCTTA GATGGGAGAT CTCAGATTTC AACACCAGCT 180  
 TTTGCAAAA CTTTCGATG CCCACCAAGC TTACCTAAAG CTACTAGAAA GGCCTTGGGA 240  
 ACTGTCAACA GAGCTACAGA AAGTCTGTA AAGACCAAGG GACCCCTCAA ACAAAAACAG 300  
 CCAAGCTTTT CTGCCAAAA GATGACTGAG AAGACTGTTA AAGCAAAAAG CTCTGTTCTT 360  
 50 GCCTCAGATG ATGCCTATCC AGAAATAGAA AAATCTTTC CCTTCAATCC TCTAGACTTT 420  
 GAGAGTTTTG ACCTGCCTGA AGAGCACCAG ATTGCGCACC TCCCCTTGAG TGGAGTGCCT 480  
 CTCATGATCC TTGACGAGGA GAGAGAGCTT GAAAAGCTGT TTCAGCTGGG CCCCCTTCA 540  
 CCTGTGAAGA TGCCCTCTCC ACCATGGGAA TCCTAATCTGT TGCACTCTCC TTCAAGCATT 600  
 CTGTGAGACC TGGATGTTGA ATTGCCACCT GTTTGCTGTG ACATAGATAT TTAATTTCT 660  
 55 TAGTCTCTCA GAGTTTGTGT GTATTGTAT TAATAAGACA TTCTTCAACA GAAAAAATA 720  
 AAAAAAA

Seq ID NO: 367 Protein sequence

Protein Accession #: NP\_004210

1 11 21 31 41 51  
 MATLIYVDKE NGEPGTRVVA KDGLKLGSGP SIKALDGRSQ VSTPRFGKTF DAPPALPKAT 60  
 65 RKALGTVNRA TEKSVKTKGP LKQKQPSFSA KRMTEKTVKA KSSVPASDDA YPEIEKFPPF 120  
 NPLDFESFDL PEEHQIAHLP LSGVPLMILD EERELEKLFQ LGPPSPVKMP SPPWESNLLQ 180  
 SPSSILSTLD VELPPVCCDI DI

Seq ID NO: 368 DNA sequence

Nucleic Acid Accession #: NM\_000597

Coding sequence: 118-1104

1 11 21 31 41 51  
 75 ATTCCGGGCG AGGAGGAGG AAGAAGCGGA GGAGGCGGCT CCGCTCGCA GGGCCGTGCA 60  
 CTGCCCCGCC CGCCGCTCG CTGCTCGCC CGCCGCGCG CGCTGCGSAC CGCCAGCATG 120  
 CTGCGGAGAG TGGGCTGCC CGCGCTGCC CTGCGCCGC CGCGCTGCT GCGCTGCTG 180  
 CGCGTGTGCT TGCTGCTACT GGGCGCGAGT GGGCGCGCG GCGGGGCGCG CGCGGAGGTG 240  
 CTGTTCCGCT GCGCGCTCG CACACCGAG CGCTGCGCG CCGTGGCGG CCGCGCGGTT 300  
 80 GCGCGCGCG CGCGGCTGCG CGCAGTGGCC GGAGGCGGCC GCATGCCATG CGCGGAGCTC 360  
 GTCCGGGAGC CGGCTGCGG CTGCTGCTCG GTGTCGCGC GGCTGAGAG CGAGGCGTGC 420  
 GGCGTCTACA CCGCGCGCTG CGGCGAGGG CTGCGCTGCT ATCCCCACCC GGGCTCCGAG 480  
 CTGCCCCCTG AGGCGCTGCT CATGGGCGAG GGCATTGTG AGAAGCGCG GAGCGCGAG 540  
 TATGCGGCA GCGCGGAGCA GGTTCAGAC AATGGCGAT ACCACTCAGA AGGAGGCGCT 600  
 85 GTGAGAAACC ACGTGGACAG CACCATGAAC ATGTTGGGCG GGGGAGGCG TGCTGGCCGG 660  
 AAGCCCTCA AGTGGGTAT GAAGGAGCTG GCGGTGTTCC GGGAGAAAGT CACTGAGCAG 720  
 CACCGGCGAG TGGGCAAGGG TGGCAAGCAT CACCTTGGCC TGGAGGAGCC CAAGAAGCTG 780  
 CGACCAACCC CTGCCAGGAC TCCCTGCCAA CAGGAATGCG ACCAGGTCTT GGAGCGGATC 840

## WO 02/086443

PCT/US02/12476

5  
10  
TCCACCATGC GCCTTCCGGA TGAGCGGGGC CCTCTGGAGC ACCTCTACTC CCTGCACATC 900  
CCCAACTGTG ACAAGCATGG CCTGTACAAC CTCAAACAGT GCAAGATGTC TCTGAACGGG 960  
CAGCGTGGGG AGTGTCTGGT TGTGAACCCC AACACCCGGA AGCTGATCCA GGGAGCCCCC 1020  
ACCATCCGGG GGGACCCCGA GTGTCTATCT TTCTACAATG AGCAGCAGGA GGCCTGCGGG 1080  
GTGCACACCC AGCGGATGCA GTAGACCGCA GCCAGCGGCT GCCTGGCGCC CCTGCCCCCC 1140  
GCCCTCTCC AACACCCGCG AGAAAACGGA GAGTGCTTGG GTGGTGGGTG CTGAGGATT 1200  
TTCCAGTTCT GACACACGTA TTTATATTTG GAAAGAGACC AGCACCAGAC TCGSCACCTC 1260  
CCCAGCTCT CTCTTCCAG CTGCAGATGC CACACCTGCT CCTTCTTGCT TTCCCCGGGG 1320  
GAGGAAGGGG GTTGTGGTCG GGGAGCTGGG GTACAGGTTT GGGGAGGGGG AAGAGAAATT 1380  
TTTATTTTGG AACCCCTGTG TCCCTTTTGC ATAAGATTAA AGGAAGGAAA AGT

Seq ID NO: 369 Protein sequence

Protein Accession #: NP\_000588

15  
20  
25  
1 11 21 31 41 51  
MLPRVGCPL FLPPPLLLPL LPLLLLLLGA SGGGGGARAE VLFRCPPCTP ERLAACGPPP 60  
VAPPAVAVA AGGARMPCAE LVREPGCGCC SVCARLEGEA CGVYTPROGQ GLRCYPHPGS 120  
ELPLQALVMG EGTCEKRRDA EYGASPEQVA DNGDDHSEGG LVENHVDSTM NMLGGGGSAG 180  
RKPLKSGMKE LAVFREKVTE QHRQMGKGGK HRLGLEBEPK LRPPPARTPC QQLDQVLER 240  
ISTMRLEPDER GPLEHLYSLH IPNCDKHGLY NLKQCKMSLN GQRGECWCVN PNTGKLIQGA 300  
FTIRGDEPECH LFYNEQOEAC GVHTQRMQ

Seq ID NO: 370 DNA sequence

Nucleic Acid Accession #: NM\_004264

Coding sequence: 6-440

30  
35  
40  
1 11 21 31 41 51  
GGAACATGGC GGATCGGCTC ACGCAGCTTC AGGACGCTGT GAATTCGCTT GCAGATCAGT 60  
TTTGTAAATGC CATTGGAGTA TTGCAGCAAT GTGGTCCCTCC TGCTCTTTTC AATAATATTC 120  
AGACAGCAAT TAACAAAGAC CAGCCAGCTA ACCCTACAGA AGAGTATGCC CAGCTTTTTC 180  
CAGCATGAT TGACAGAAC GCAAAAGACA TTGATGTTTT GATAGATTCC TTACCCAGTG 240  
AAGAATCTAC AGCTGCTTTC CAGGCTGCTA GCTGTATATA GCTAGAAGAA GAAACCATG 300  
AAGCTGCTAC ATGTGTGGAG GATGTTGTTT ATCGAGGAGA CATGCTTCTG GAGAAGATAC 360  
AAAGCGCACT TGCTGATATT GCACAGTCAC AGCTGAAGAC AAGAAGTGGT ACCCATAGCC 420  
AGTCTCTTCC AGACTCATAG CATCAGTGGG TACCATGTGG CTGAGAAAAG AACTGTTTGA 480  
GTGCCATTAA GAATCTCTCA TCAGACTTAG ATACAAGCCT TACCAACAAT TACAGAAACA 540  
TTAAACACTA TGACACATTA CCTTTTAGC TATTTTAAAT AGTCTTCTAT TTTCACTCTT 600  
GATAAGCTTA TAAATCATGA TTGAATCAGC TTAAAGCAT CATACCATCA TTTTTTAACT 660  
GAGTGAATTT ATTAAGGCAT GTAATACATT AATGAACATA ATATAAGGAA ACATATGTAA 720  
AATTCGTGTA TGACATAATT TAATCTCCA TTTTGTGTA TTGGCCAGTA CTTTACAAT 780  
C

Seq ID NO: 371 Protein sequence

Protein Accession #: NP\_004255

50  
55  
1 11 21 31 41 51  
MADRLTQLQD AVNSLADQFC NAIGVLQCG PPASPNNIQT AINKDQPANP TEEYAQLFAA 60  
LIARTAKDID VLIDSLPSEB STAALQAASL YKLEENHEA ATCVEDVVYR GDMLEKIQS 120  
ALADIAQSQL KTRSGTHSQS LPDS

Seq ID NO: 372 DNA sequence

Nucleic Acid Accession #: AJ271091

Coding sequence: 1-1113

60  
65  
70  
75  
80  
1 11 21 31 41 51  
ATGGAGAATC AGTGTGTCG GCGGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60  
CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120  
CATTTCAAAG CTCAGGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180  
TTCTTAGACC TTGTGAAACC AGAGCCTGTT TACAAACTGA CCCAGAGGCA GGTAAACATT 240  
ACAGTACAGA AGAAAGTGGT TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300  
CTGTTTTTGG CTCCTGACTT TGATGTTGG CTGGATGAAT CTGATGGGGA AATGGAGCTC 360  
AGAGCTAAGG AAGAAGACCG CCTAAATAAA CTCCGACTGG AAAGCGAAGG CTCCTCTGAA 420  
ACTCTTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTGCA ATTCTTGGGA 480  
TTCTCCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCCTTT 540  
TATGACACAT TCCATCTGTT GGCTGACATG ATGTATTTCT GCCAGATGCT GGCAGTTGTG 600  
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACGTCACCGG TGCTGCCTTC TCTGATCCAG 660  
CTTCTTGGAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGAGAAC 720  
AAAGCTGTGG TTTTCTTGT GTTTTATTG TGGAGTGCAA TTGAAATTTT CAGTACTCT 780  
TTCTACATGC TGACGTGCAT TGACATGGAT TGGAGGTGTC TCACATGGCT TCGTTACACT 840  
CTGTGGATTC CCTATATCC ACTGGGATGT TTGGGGAAG CTGTCTCAGT GATTCACTCC 900  
ATTCCAATAT TCAATGAGAC CGACGATTC AGTTTCACAT TGCCATATCC AGTGAATATC 960  
AAAGTTAGAT TTTCTTTTTC TCTTCAGATT TATCTTATAA TGATATTTT AGGTTTATAC 1020  
ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGCTT 1080  
CATGCCTGTG ATCCACGCGC TTTGGGAGGC TGA

Seq ID NO: 373 Protein sequence

Protein Accession #: CAB69070

85  
1 11 21 31 41 51  
MENQVLTFHV YNAQRHRELY LRVELSDVON PAISITENVL HFKAQGHGAK GDNVYEFHLE 60  
FLDLVKPEPV YLTKRQVNI TVQKKVSQWN ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120  
RAKEBERLNK LRLSEGSPE TLTNLRKGYL FMYNLVQFLG FSWIFVNLTV RFLCLGKESF 180

YDTFHTVADM MYFQMLAVV ETINAAIGVT TSPVLP SLIQ LLGRNFILFI IFGTMBEQN 240  
KAVVFPVYFL WSAIBIFRYS FYMLTICIDMD WKVLTLWLYT LWIPLYPLGC LAEAVSVIQS 300  
IPIFNETGRF SFTLPYPVKI KVRFSFPLQI YLIMIFLGLY INFRELYKQR RLKMRAGAVA 360  
HACDPSALGG

Seq ID NO: 374 DNA sequence  
Nucleic Acid Accession #: NM\_016395  
Coding sequence: 1-1113

1 11 21 31 41 51  
| | | | | |  
ATGAGAGATC AGGTGTTGAC GCCGCATGTC TACTGGGCTC AGCGACACCG CGAGCTATAT 60  
CTGCGCGTGG AGCTGAGTGA CGTACAGAAC CCTGCCATCA GCATCACTGA AAACGTGCTG 120  
15 CATTTCAAAG CTCAGAGACA TGGTGCCAAA GGAGACAATG TCTATGAATT TCACCTGGAG 180  
TCTCTAGACC TTGTGAAACC AGAGCCTGTT TACAACTGA CCCAGAGGCA GGTAAACATT 240  
ACAGTACAGA AGAAAGTGAG TCAGTGGTGG GAGAGACTCA CAAAGCAGGA AAAGCGACCA 300  
CTGTTTTTGG CTCCTGACTT TGATCGTTGG CTGGATGAAT CTGATGCGGA AATGGAGCTC 360  
AGAGCTAAGG AAGAAGAGCG CCTAAATAAA CTCGACTGG AAAGCGAAGG CTCTCTGAA 420  
20 ACTCTACAA ACTTAAGGAA AGGATACCTG TTTATGTATA ATCTTGTCGA ATCTTGCGGA 480  
TCTCTCTGGA TCTTTGTCAA CCTGACTGTG CGATTCTGTA TCTTGGGAAA AGAGTCTTTT 540  
TATGACACAT TCCATCTGTT GGCTGACATG ATGTATTTCT GCCAGATGCT GGCAGTTGTG 600  
GAAACTATCA ATGCAGCAAT TGGAGTCACT ACCTCACCGG TGCTGCTTTC TCTGATCCAG 660  
CTTCTTGGAA GAAATTTTAT TTTGTTTATC ATCTTTGGCA CCATGGAAGA AATGCAGAAC 720  
25 AAAGCTGTGG TTTTCTTTGT GTTTTATTTG TGGAGTGCAA TTGAAATTTT CAGGTACTCT 780  
TTCTACATCG TGAAGTGATG TGACATGGAT TGGAGGTGCT TCACATGGCT TCGTTACACT 840  
CTGTGGATTC CCTATATACC ACTGGGATGT TTGGCGGAAG CTGTCTCAGT GATTTCAGTC 900  
ATTCCAATAT TCAATGAGAC CGGACGATTC AGTTTCACAT TGCCATATCC AGTGAAATC 960  
30 AAAGTTAGAT TTTCTTTTTC TCTTCAGATT TATCTTATAA TGATATTTTT AGGTTTATAC 1020  
ATAAATTTTC GTCACTTTTA TAAACAGCGC AGACTGAAAA TGAGGGCAGG CGCAGTGGCT 1080  
CATGCTGTG ATCCACAGCGC TTTGGGAGGC TGA

Seq ID NO: 375 Protein sequence  
Protein Accession #: NP\_057479

1 11 21 31 41 51  
| | | | | |  
MENQVLTPHV YWAQRHRELY LRVELSDVQN PAISITENVL HFKAQGHGAK GDNVYEFHLE 60  
40 FLDLVKPEPV YKLTQRQVNI TVQKRVSQNW ERLTKQEKRP LFLAPDFDRW LDESDAEMEL 120  
RAKEEERLNK LRLESEGSPE TLTLNRKGYL FMYNLVQFLG PSWIFVNLTV RFCILGKESF 180  
YDTFHTVADM MYFQMLAVV ETINAAIGVT TSPVLP SLIQ LLGRNFILFI IFGTMBEQN 240  
KAVVFPVYFL WSAIBIFRYS FYMLTICIDMD WKVLTLWLYT LWIPLYPLGC LVEAVSVIQS 300  
45 IPIFNETGRF SFTLPYPVKI KVRFSFPLQI YLIMIFLGLY INFRELYKQR RRRYKGRKR 360  
STKKRDLDF LPV

Seq ID NO: 376 DNA sequence  
Nucleic Acid Accession #: NM\_005987  
Coding sequence: 1-270

1 11 21 31 41 51  
| | | | | |  
ATGAATTCTC AGCAGCAGAA GCAGCCTTGC ACCCCACCCC CTCAGCCTCA GCAGCAGCAG 60  
55 GTGAAACAAC CTGCGCAGCC TCCACCCAGC GAACCATGCA TCCCAAAAC CAAGGAGCCC 120  
TGCCAAACCA AGGTGCTGTA GCCCTGCCAC CCCAAAGTGC CTGAGCCCTG CCAGCCCAAG 180  
ATTCCAGAGC CCTGCCAGCC CAAGTGCCT GAGCCCTGCC CTTCAACGGT CACTCCAGCA 240  
CCAGCCAGC AGAAGACCAA GCAGAAGTAA

Seq ID NO: 377 Protein sequence  
Protein Accession #: NP\_005978

1 11 21 31 41 51  
| | | | | |  
MNSQQQKQPC TPPPQQQQQ VRQPCQPPQ EPCIPKTEP CQPKVPEPC PKVPEQPPK 60  
65 IPEPCQPKVP EPCPSTVTPA PAQQTQKQK

Seq ID NO: 378 DNA sequence  
Nucleic Acid Accession #: NM\_002105  
Coding sequence: 74-505

1 11 21 31 41 51  
| | | | | |  
75 ACAGCAGTTA CACTGCGGCG GGCGTCTGTT CTAGTGTITG AGCCGTCGTG CTTACCCGGT 60  
CTACCTCGCT AGCATGTCGG GCGGCGGCAA GACTGGCGGC AAGGCCCGCG CCAAGGCCAA 120  
GTGCGGCTCG TCGCGCGCGG GCCTCCAGTT CCCAGTGGGC CGTGTACACC GGCTGCTGCG 180  
GAAGGCCACG TAGCGCGAGC GCGTTGGGCG CGGCGCGCCA GTGTACCTGG CCGCAGTGCT 240  
GGAGTACCTC ACCGCTGAGA TCCTGGAGCT GCGGGGCAAT GCGGCCGCGG ACAACAAGAA 300  
GAGCGGAATC ATCCCCGCC ACTCGAGCT GGCCATCCGC AACGACGAGG AGCTCAACAA 360  
80 GCTGCTGGGC GCGGTGACGA TCGCCAGGG AGGCGTCTCT CCCAACATCC AGGCGGTGCT 420  
GCTGCCCAAG AAGACGAGCG CCACCGTGGG GCGAAGGCG CCTCGGGCG GCAAGAAGGC 480  
CAGCCAGGCC TCCAGGAGT ACTAAGAGG CCGCGCGCGG GCGCGCGCGC CCGAGCTCCC 540  
CATGCCACCA CAAAGGCCCT TTTAAGGGCC ACCACGCGCC TCATGGAAG AGCTGAGCCG 600  
CTTCAGACTG CGGGGCAAGC GGGCGCGGCG TCCCTTCCCC TCCCTCCCC TCGCCCGCT 660  
85 TCGCGCGCGG GCGTGAATC CCGCGCGGCC CCGCTCCCG TCCCGCACCG CCGCGCGGT 720  
CGGCCTCGGG CCGTCCCTGT CCGCGCTCCG CCTCCGGTA GGGTTCGGGC CTTCCGATG 780  
CGGCTTGGGC GCTCTCTCGG GACCTCCGTG GCGCGGAAGA CCGGAGCCTG CCGGGGGAG 840

GCCGCGCGCG COGCACTGCG CGGCGCTCGG GTTCGTGACT CAGCCGCCCC ATCCCCGAGTC 900  
 GCTAAGGGGC TGGGGGGAGG CGGAGCAGCC TTCTGGAAGA CTGGGCTTC CGCTCTGAGC 960  
 CAGGGCGGAG GTGGGCGAGT CAGGCGGAGA GCGGCGCGCC CTGAAGGTGA GTGAGGCCCT 1020  
 CGGAGCTGCG AGCGGGGGTG TCTGGTACCC CCGGCGCGTG GTGCTTAGCC CAGGACTTTC 1080  
 AGACGGCGCG TGGCGGGGAG GCTTTGGTGG GAGAGACGCG ATCGCGGATT TCGGTCTGGC 1140  
 GCGGCTTCTG CGGCGGGGAG CCAGGCGCTT CACATCAGCT CTCCTCCAT CTTCACTCAT 1200  
 AGGTCTGCGC TGGGGCGGAG ACGAAGCACT TGGTAACAGG CACATCTTCC TCCCGAGTGA 1260  
 CTGCTCTCTA GGAGGACATT TAGGGGAGGG CAGAGGCGCT CAGTTTGGCT TCACGGCTGG 1320  
 CTATGTGAGC AGCAAGAGTC GTTTTGGGGA ACGGAGCTGG CAGCCAGGCC TGTGCGGGCC 1380  
 CGAGCGCGCG CCAATTTCCC TTCCAGCAAA CTCAACTCGG CAATCCAAGC ACCTAGATAC 1440  
 CAGCACAAGT CGGTTAATCC CTGTCTGGAC TGAGCCTCGG TTGGCTTCTG AACTGGAATT 1500  
 CTGACAGTAA CCGTTCCAGC ACTAGAACCT TAGGCATTGG GGAGTTTATG ATGGACTAAT 1560  
 TTTATTAAAG GATTGTTTTT TTTT

Seq ID NO: 379 Protein sequence  
 Protein Accession #: NP\_002096

1 11 21 31 41 51  
 | | | | |  
 MSGRGKTGK ARAKAKSRSS RAGLQFPVGR VHRLLRKHGY AERVGAGAPV YLAAVLEYLT 60  
 AEILELAGNA ARDNKTRII PRHLQLAIRN DEELNKLGG VTIAGGGVLP NIQAVLLPKK 120  
 TSATVGPAP SGGKKATQAS QEY

Seq ID NO: 380 DNA sequence  
 Nucleic Acid Accession #: AL136942  
 Coding sequence: 184-864

1 11 21 31 41 51  
 | | | | |  
 ACGGTCGCGG CAGAAGCTCG GAGCTCTCGG GGTATCGAGG AGGCAGGCC GCGGGCGCAC 60  
 GGGCGAGCGG GCGGGGAGCC GGAGCGCGCG AGGAGCGGCG AGCAGCGGCG CCGCGGGCTC 120  
 CAGGCGAGCG GGTCCAGCGT CCTGAAACT TGCGCGCGCG CTCGCGCCAC TGCGCCCGGA 180  
 GCGATGAAGA TGGTCGCGCC CTGGACGCGG TTCTACTCCA ACAGCTGCTG CTTGTGCTGC 240  
 CATGTCCGCA CCGGCACCAT CTGTCTCGGC GTCTGGTATC TGATCATCAA TGCTGTGTTA 300  
 CTGTTGATTT TATTGAGTGC CCGGCTGATG COGATCAGT ATAACCTTTC AAGTCTGAA 360  
 CTGGGAGGTG ACTTTGAGTT CATGGATGAT GCCAATCTGT GCATTCGCAT TGGGATTCTT 420  
 CTCTCATGTA TCGTATATG TGCTATGGCT ACTTACGAG CGTACAAGCA ACGCGCAGCC 480  
 TGGATCATCC CATTCTTCTG TTACCATGAT TTTGACITTT CCCTGAACAT GTTGGTTGCA 540  
 ATCATGTGTC TTATTTATCC AAACCTCCAT CAGGAATACA TACGGCAACT GCCTCCTAAT 600  
 TTTCCTTACA GAGATGATGT CATGTGAGTG AATCCTACCT GTTTGGTCTT TATTATTCTT 660  
 CTGTTTATTA GCATTATCTT GACTTTTAA GGTACTTGA TTGCTGTGT TTGGAATGTC 720  
 TACCGATACA TCAATGGTAG GAATCCTCT GATGTCTGG TTTATGTTAC CAGCAATGAC 780  
 ACTACGCTGC TGCTACCCCG GTATGATGAT GCCACTGTGA ATGGTGTCTG CAAGAGGCCA 840  
 CCGCCACCTT ACGTGTCTGC CTAAGCCTTC AAGTGGGCGG AGCTGAGGCG AGCAGCTTGA 900  
 CTTGTCAGAC ATTGAGCAA TAGTTCTGTT ATTTCACTTT TGCCATGAGC CTCTCTGAGC 960  
 TTGTTTGTG CTGAAATGCT ACTTTTAAA ATTTAGATGT TAGATTGAAA ACTGTAGTTT 1020  
 TCAACATATG CTTTGTCTGA ACACGTGATG AGATTAACCT TAGAATCTCT CCGTACGAT 1080  
 TGGGATATA ACGGGCTTCA CTAACCTTCC CTAGGCATTG AAACCTCCCC CAAATCTGAT 1140  
 GGACCTAGAA GTCTGCTTTT GTACCTGCTG GCGCCCAAGG TTGGGCATTT TTCTCTCTGT 1200  
 TCCTCTCTT TTGAAATGT AAAATAAAG CAAAATAGA CAACCTTTTC TTCAGCCATT 1260  
 CCAGCATAGA GAACAAACCT TTATGGAAAC AGGAATGTCA ATTTGTATAT CATTGTTCTA 1320  
 ATTAGGTAAA TAGAATGCTT TATGTATGTG TTACAAGAAT TTCCCCCACA ACATCCTTTA 1380  
 TGACTGAAGT TCAATGACG TTTGTGTTG GTGGTAAAGG ATTTCTCTCA TGGCCTGAAT 1440  
 TAAGACCAT AGAAGACACC AGGCGGTGGG AGCAGTGACC ATCTACTGAC TGTCTCTGTG 1500  
 GATCTTGTGT CCAGGACAT GGGGTGACAT GCCTGTATG TGTAGAGGG TGAATGAT 1560  
 GTGTTTGGCG CTGCATGGGA TCTGTGCCC CTCTCTCTCT GGATTCACAT CCCCACCGAG 1620  
 GGGCGGCTTT TACTACCTAG TCTGCCCTAG ATTGGTTCAA GGAGGTGATC CAACCTGACT 1680  
 TATCAAGTGG AATTGGGATA TATTGTATAT ACTTCTGCTT AACACATGG AAAAGGTTT 1740  
 TCTTTTCCCT GCAAGCTACA TCCTACTGCT TTGAACCTCC AAGTATGTCT AGTCACCTTT 1800  
 TAAATGTAA ACATTTTCAG AAAAATGAGG ATTGCCCTCC TTGTATGCGC TTTTACCTT 1860  
 GACTACCTGA ATTGCAAGGG ATTTTATAT ATTCATATGT TACAAAGTCA GCAACTCTCC 1920  
 TGTGTTTCA TTATTGAATG TGCTGTAAAT TAAGTCGTTT GCAATTAATA CAAGGTTTTC 1980  
 CCACATCCAA AAAAAAAAAA AAAAA

Seq ID NO: 381 Protein sequence  
 Protein Accession #: CAB66876

1 11 21 31 41 51  
 | | | | |  
 MKMVAPWTRF YSNSCLCCH VRTGTILLGV WYLIINAVVL LILLSALADP DOYNFSSSEL 60  
 GGDPEFMDA NMCIAlAISL LMILICAMAT YGAYKQRAAW IIPFFCYQIF DFLNMLVAI 120  
 TVLIYPNSIQ EYIRQLPPNF PYRDDVMSVN PTCLVLIILL FISIIITFKG YLISCVWNCY 180  
 RYINGRNSD VLVVYTSNDT TVLLPPYDDA TVNGAAKEPP PPVYSA

Seq ID NO: 382 DNA sequence  
 Nucleic Acid Accession #: NM\_002510  
 Coding sequence: 92-1774

1 11 21 31 41 51  
 | | | | |  
 CAGATGCCAG AAGAACACTG TTGCTCTTGG TGGACGGGCC CAGAGGAATT CAGAGTTAAA 60  
 CCTGAGTGC CTGCGTCCGT GAGAATTCAG CATGGAATGT CTCTACTATT TCCTGGGATT 120  
 TCTGCTCCTG GCTGCAAGAT TGCCACTTGA TGCCGCCAAA CGATTTCATG ATGTGCTGGG 180  
 CRATGAAGA CCTCTCTGCT ACATGAGGGA GCACAATCAA TTAATGGCT GGTCTCTGA 240  
 TGAATGAC TGAATGAAA AACTCTACCC AGTGTGGAAG CGGGGAGACA TGAGGTGAA 300  
 AACTCCTGG AAGGGAGGCC GTGTGAGGC GGTCTGACC AGTGACTCAC CAGCCCTCGT 360

	GGGCTCAAAAT	ATAACATTTG	CGGTGAACCT	GATATTCCTT	AGATGCCAAA	AGGAAGATGC	420
	CAATGGCAAC	ATAGTCTATG	AGAAGAACTG	CAGAAATGAG	GCTGGTTTAT	CTGCTGATCC	480
	ATATGTTTAC	AACCTGGACAG	CATGGTCAGA	GGACAGTGAC	GGGGAAAATG	GCACCGGCCA	540
5	AAGCCATCAT	AACCTCTTCC	CTGATGGGAA	ACCTTTTCTT	CACCACCCCG	GATGGAGAG	600
	ATGGAATTTT	ATCTAGTCTT	TCCACACACT	TGGTCAGTAT	TTCCAGAAAT	TGGGACGATG	660
	TTCACTGAGA	GTTTCTGTGA	ACACAGCCAA	TGTGACACTT	GGGCTCAAC	TCATGGAAGT	720
	GACTGTCTAC	AGAAGACATG	GACGGGCATA	TGTTCCCATC	GCACAAGTGA	AAGATGTGTA	780
	CGTGGTAACA	GATCAGATTG	CTGTGTTTGT	GACTATGTTC	CAGAAGAAGC	ATCGAAATTC	840
10	ATCCGACGAA	ACCTTCTCTA	AAGATCTCCC	CATTATGTTT	GATGTCTCTGA	TTCTATGATCC	900
	TAGCCACTTC	CTCAATTTAT	CTACCATTA	CTACAAGTGG	AGCTTGGGG	ATAATACTGG	960
	CCTGTTTGGT	TCCACCAATC	ATACTGTGAA	TCACAAGTAT	GTGCTCAATG	GAACCTTCAG	1020
	CCTTAACCTC	ACTGTGAAG	CTGCAGCACC	AGGACCTTGT	CCGCCACCGC	CACCACCACC	1080
	CAGACCTTCA	AAACCCACCC	CTTCTTTAGG	ACCTGTCTGT	GACAACCCCC	TGGAGCTGAG	1140
15	TAGGATTCTT	GATGAAAACT	GCCAGATTAA	CAGATATGGC	CACTTTCAAG	CCACCATCAC	1200
	AATTGTAGAG	GGAACTCTAG	AGGTTAAACAT	CATCCAGATG	ACAGACGTCC	TGATGCGGTT	1260
	GCCATGGCTC	GAAAGCTCCC	TAATAGACTT	TGTCGTGACC	TGCCAAGGGA	GCATTCCCAT	1320
	GGAGGTCTGT	ACCATCATTT	CTGACCCAC	CTGCGAGATC	ACCCAGAAAC	CAGTCTGCAG	1380
	CCCTGTGGAT	GTGGATGAGA	TGTGTCTGCT	GACTGTGAGA	CGAACCTTCA	ATGGGCTGCG	1440
20	GAOCTACTGT	GTGAACCTCA	CCCTGGGGGA	TGACACAAGC	CTGGCTCTCA	CGAGCACCTC	1500
	GATTTCTGTT	CCTGACAGAG	ACCCAGCCTC	GCCTTTAAGG	ATGGCAAAAC	GTGCCCTGAT	1560
	CTCGTGTGGC	TGCTGGGACA	TATTTGTGAC	TGTGATCTCC	CTCTGTGGT	ACAAAAACAA	1620
	CAAGGAATAC	AACCCATATG	AAAATAGTCC	TGGGAATGTG	GTCAAGAGCA	AAGGCTCGAG	1680
	TGTCTTTCTC	AACCGTGCAA	AAGCCGTGTT	CTTCCCGGGA	AACCAGGAAA	AGGATCCGCT	1740
25	ACTCAAAAC	CAGAATTTTA	AAGGAGTTTC	TAAATTTTCG	ACCTTGTGTT	TGAAGCTCAC	1800
	TTTTCAGTGC	CATTGATGTG	AGATGTGCTG	GAGTGGCTAT	TAACTTTTTC	TTCTTAAAGA	1860
	TTATTTGTTA	ATAGATATTG	TGGTTTGGGG	AAGTTGAATT	TTTTATAGTT	TAAATGTCTAT	1920
	TTTAGAGATG	GGGAGAGGGA	TTATATGCA	GGCAGCTTCA	GCCATGTTGT	GAAACTGATA	1980
	AAAGCAACTT	AGCAAGGCTT	CTTTTCATTA	TTTTTTATGT	TTCACTTATA	AAGTCTTAGG	2040
30	TAAGTAGTAG	GATAGAAACA	CTGTGTCCCG	AGAGTAAGGA	GAGAAGCTAC	TATTGATTAG	2100
	AGCCTAACCC	AGGTAACTG	CAAGAAGAGG	CGGGATACCT	TCAGCTTTCC	ATGTAAGTGT	2160
	ATGCATAAAG	CCATGATAGT	CCAGTTTCTA	AGATCATGTT	CCAAGCTAAC	TGAATCCAC	2220
	TTCAATACAC	ACTCATGAAC	TCCTGATGGA	ACAAATACAG	GCCCAAGCCT	TGGGTATGAT	2280
	GTGCACACTT	GCTAGACTCA	GAAAAAATAC	TACTCTCATA	AATGGGTGGG	AGTATTTTGG	2340
35	TGACAACTTA	CTTTGCTTGG	CTGAGTGAAG	GAATGATATT	CATATATTCA	TTTATTCAT	2400
	GGACATTTAG	TTAGTGCTTT	TTATATACCA	GGCATGATGC	TGAGTGACAC	TCCTGTGAT	2460
	ATTTCCAAAT	TTTGTATAG	TGCTGACACA	TATTTGAAAT	CATATATTAA	GACTTTCCAA	2520
	AGATGAGGTC	CCTGGTTTTT	CATGGCAACT	TGATCAGTAA	GGATTTCCAC	TCTGTTTGTA	2580
40	ACTAAAACCA	TCTACTATAT	GTTAGACATG	ACATTTCTTT	TCTCTCTCTT	CTGAAAAATA	2640
	AAGTGTGGGA	AGAGACAAAA	AAAAAATA				

Seq ID NO: 383 Protein sequence  
Protein Accession #: NP\_002501

	1	11	21	31	41	51	
45	MECLYYFLGF	LLLAARLPD	AAKRFHDVLG	NERPSAYMRE	HNQLNGWSSD	ENDWNEKLYP	60
	VWKRGRDMRK	NSWKGGRVQA	VLTSDSPALV	GSNITFAVNL	IPPRQKEDA	NGNIVYEKNC	120
	RNEAGLSADP	YVYNWTANSE	DSDGENGTCG	SHENVFPDGG	PPPHPGWR	WNFIYVFHTL	180
50	QGYFQKLGR	SVRVSVNTAN	VTLPQLMEV	TVYRRHGRAY	VPIAQKDVY	VVTDIQPFV	240
	TMFQKNDNR	SDETFLKDLF	IMFDVLIHDP	SEFLNYSTIN	YKWSFGDNTG	LPVSTNHTVN	300
	HTYVLNGTFS	LNLTVKAAAP	GPCPPPPPPP	RPSKPTPSLG	PAGDNPLELS	RIPDENCQIN	360
	RYGHFQATIT	IVGSIILEVNI	IQMTDVLMPV	PWPESSLIDF	VVTQGSIPT	EVCTIISDPT	420
	CEITONTVC	PVDVDEMCLL	TVRRTPNGSG	TYCVNLTLGD	DTSLALTSTL	ISVPRDRPAS	480
55	PLRMANSALI	SVGCLAIPTV	VISLLVYKKE	KEYNPIENSP	GNVVRSGKLS	VFLNRKAVF	540
	PPGNQEKDPL	LKNQEPKQVS					

Seq ID NO: 384 DNA sequence  
Nucleic Acid Accession #: NM\_001134  
Coding sequence: 48-1877

	1	11	21	31	41	51	
60	TCCATATTGT	GCTTCCACCA	CTGCCAATAA	CAAAATAACT	AGCAACCATG	AAGTGGGTGG	60
	AATCAATTTT	TTTAATTTTC	CTACTAAAT	TTACTGAATC	CAGAACACTG	CATAGAAATG	120
65	AATATGGAAT	AGCTTCCATA	TTGGATTCTT	ACCAATGTAC	TGCAGAGATA	AGTTTAGCTG	180
	ACCTGGCTAC	CATATTTTTT	GCCCAAGTTG	TTCAAGAAGC	CACCTTACAG	GAGTAAGCA	240
	AAATGGTGAA	AGATGCATTG	ACTGCAATTG	AGAAACCCAC	TGGAGATGAA	CAGTCTTCAG	300
	GGTGTTTAGA	AAACCAAGTA	CCTGCCTTTC	TGGAAGAACT	TTGCCATGAG	AAAGAAATTT	360
70	TGGAGAAGTA	CGGACATTCA	GACTGTGCA	GCCAAAGTGA	AGAGGGAAGA	CATAACTGTT	420
	TTCTTGACCA	CAAAAAGCCC	ACTCCAGCAT	CGATCCCACT	TTTCCAAGTT	CCAGAACCTG	480
	TCACAAGCTG	TGAAGCATAT	GAAGAAGACA	GGGAGACATT	CATGAACAAA	TTCAATTTATG	540
	AGATAGCAAG	AAGGCATCCC	TTCTGTATG	CACCTACAAT	TCTTCTTTGG	GCTGCTCGCT	600
	ATGACAAAAT	AATTCCATCT	TGCTGCAAG	CTGAAAATGC	AGTTGAATGC	TTCCAAACAA	660
	AGGCAGCAAC	AGTTACAAAA	GAATTAAGAG	AAAGCAGCTT	GTTAAATCAA	CATGCATGTG	720
75	CAGTAATGAA	AAATTTTGGG	ACCGAACTT	TCCAAGCCAT	AACGTGTTACT	AACTGAGTCT	780
	AGAAGTTTAC	CAAAGTTAAT	TTTACTGAAA	TCCAGAAACT	AGTCTGGAT	GTGGCCCATG	840
	TACATGAGCA	CTGTGTCAGA	GGAGATGTGC	TGGATTGTCT	GCAGGATGGG	GA AAAAATCA	900
	TGTCTTACAT	ATGTTCTCAA	CAAGACACTC	TGTCAAACAA	AATAACAGAA	TGCTGCAAAAC	960
80	TGACCCAGCT	GGAAAGTGTG	CAATGTATAA	TTATGTCAGA	AAATGATGAA	AAACCTGAAG	1020
	GTCTATCTCC	AAATCTAAAC	AGGTTTITAG	GAGATAGAGA	TTTTAACCAA	TTTTCTTCAG	1080
	GGGAAAAAAA	TATCTTCTTG	GCAAGTTTITG	TTTATGAATA	TTCAAGAAGA	CATCCTCAGC	1140
	TTGCTGTCTC	AGTAATTTCTA	AGAGTTGCTA	AAGGATACCA	GGAGTTATTT	GAGAAGTGTG	1200
	TCCAGACTGA	AAACCTCTCT	GAATGCCAAG	ATAAGGAGGA	AGAAGAATTA	CAGAAATACA	1260
85	TCCAGGAGAG	CCAAGCATTTG	GCAAGCGAA	GCTGCGGCTC	CTTCCAGAAA	CTAGGAGAAAT	1320
	ATTACTTACA	AAATGCGTTT	CTGTTGCTTT	ACACAAAGAA	AGCCCCCAG	CTGACCTCTG	1380
	CGGAGCTGAT	GGCCATCAC	AGAAAAATG	CAGCCACAGC	AGCCACTTGT	TGCCAACTCA	1440
	GTGAGGACAA	ACTATTGGCC	TGTGGCGAGG	GAGCGCTGTA	CATTATTATC	GGACACTTAT	1500

GTATCAGACA TGAATGACT CCAGTAAACC CTGGTGTGG CCAGTGTGC ACTTCTTCAT 1560  
 ATGCCAACAG GAGGCCATGC TTCAGCAGCT TGGTGGTGA TGAACATAT GTCCCTCCTG 1620  
 CATCTCTGTA TGACAGTTC ATTTCCATA AGGATCTGT CCAAGCTCAG GGTGTAGCGC 1680  
 TGCAAAACGAT GAAGCAAGAG TTTCTCATT ACCTTGTGA GCAAAAGCCA CAAATAACAG 1740  
 AGGAACAAC TGAAGCTGTC ATTGCAGATT TCTCAGGCCT GTTGGAGAAA TGCTGCCAAG 1800  
 GCCAGGAACA GGAAGTCTGC TTTGCTGAAG AGGGACAAAA ACTGATTTC AAAAACTGCTG 1860  
 CTGCTTTGGG AGTTTAAATT ACTTCAGGGG AAGAGAAGAC AAAACGAGTC TTTTATTCGG 1920  
 TGTGAACATT TCTCTTTAAT TTTAACTGAT TTAACACTTT TGTGAATTA ATGAAATGAT 1980  
 AAAGACTTTT ATGTGAGATT TCCTTATCAC AGAAATAAAA TATCTCCAAA TG

Seq ID NO: 385 Protein sequence  
 Protein Accession #: NP\_001125

1 11 21 31 41 51  
 MKWVESIFLI FLNFTESRT LHRNEYGIAS ILDSYQCTAE ISLADLATIF FAQFVQEATY 60  
 KEVSKMVKDA LTAIEKPTGD EQSSGCLENQ LPAPLEELCH EKEILEKYGH SDCCSQSEEG 120  
 RHNCPLAHKK PTPASIPLFQ VPEPVTSCBA YEEDRETFFM KFIYIEARRH PFLYAPTILL 180  
 WAARYDKIIP SOCKAENAVE CFQTKAATVT KELRESSLLN QHACAVMKNF GTRTFQAITY 240  
 TKLSQKPTKV NPTIEQLKVL DVAHVHECC RGDVLDCLQD GEKIMSYICS QODTLENKIT 300  
 ECCRLTTLER GQCIIHAEND EKEPEGLSPNL NRFLGDRDFN QFSSGEKNIF LASFVHEYSR 360  
 RHPQLAVSVI LRVAKYQDEL LEKCFQTEPN LECQDKGEEB LQKYIQESQA LAKRSCGLFQ 420  
 KLGEYVLQNA FLVATYKKAP QLTSSSELMAI TRKMAATAAT CCQLSEDKLL ACCEGAADII 480  
 IGHLCIRHEM TPVNPVGQC CTSSYANRRP CFSSLVVDFT YVPPAFSDDK FIFHKDLQQA 540  
 QGVALQTMKQ EFLINLVKQK PQITEEQLBA VIADFSGLLE KCCQGGQEV CFAERGQKLI 600  
 SKTRAAALGV

Seq ID NO: 386 DNA sequence  
 Nucleic Acid Accession #: NM\_002205.1  
 Coding sequence: 1..3149

1 11 21 31 41 51  
 ATGGGGAGCC GGAAGCCAGA GTCCCTCTCT CACGCGGTGC AGCTGCGCTG GGGCCCCCGG 60  
 CGCCGACCCC CGCTSSSTGCC GCTGCTGTTC CTGCTSTGTC CGCGGCCACC CAGGGTCGGG 120  
 GGCTTCAACT TAGACGCGGA GCGCCAGACA GTACTCTCGG GCGCCCCGGG CTCTTCTCTC 180  
 GGATTCTCAG TGGAGTTTAA CCGGCGGGGA ACAGACGGGG TCAGTGTGCT GGTGGGAGCA 240  
 CCGAAGGCTA ATACCAAGCCA GCCAGGAGTG CTGCAGGGTG GTGCTGTCTA CTTCTGTCTC 300  
 TGGGGTGCCA GCGCCACACA GTGCACCCCC ATTGAATTGG ACAGCAAAAG CTCTCGGCTC 360  
 CTGGAGTCTT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCTTTCAG 420  
 TGGTTCGGGG CAACAGTTCG AGCCCATGGC TCCTCCATCT TGGCATGCGC TCCACTGTAC 480  
 AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACTGTCTA CCTCTCCACA 540  
 GATAAATTCA CCGGAATTCT GGAGATGCA CCGTCCCGCT CAGATTTCAG CTGGGCAGCA 600  
 GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT 660  
 TTAGTGGGAC CAGGAAGCTA TTTCTGGCAA GCGCAGATCC TGCTGCCAC TCAGGAGCAG 720  
 ATTGCAGAA CTTATTACCC CGAGTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT 780  
 CGCCAGGCCA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTGGTGA 840  
 TTCAGTGGTG ATGACACAGA AGACTTTGTT GCTGGTGTGC CCAAGGGGAA CCTCACTTAC 900  
 GGCTATGTCA CCATCTCTAA TGGCTCAGAC ATTCGATCCC TCTCAACTT CTCAGGGGAA 960  
 CAGATGGCTT CAGATTTGG CTATGAGTG GCGGCCACAG ACCTCAATGG GGAAGGGCTG 1020  
 GATGACTTGC TGGTGGGGGC ACCCTGTCTC ATGATCGGA CCGCTGACGG GCGGCTCAG 1080  
 GAGGTGGGCA GGGTCTACGT CTACCTGCAG CACCCAGCGG GCATAGAGCC CACGCCACCC 1140  
 CTTACCTCA CTGCCATGTA TGAGTTTGGC GATTTTGGCA GCTCTGTGAC CCGCTGGGG 1200  
 GACCTGGACC AGGATGGGTA CAATGATGTG GCCATCGGGG CTCCCTTTGG TGGGGAGACC 1260  
 CAGCAGGGAG TAGTGTTTGT ATTCTCTGGG GCGCCAGGAG GGCTGGGCTC TAAGCCTTCC 1320  
 CAGGTTCTGC AGCCCTCTGT GGCAGCCAGC CACAACCCAG ACTCTTTTGG CTCTGCCCTT 1380  
 CGAGGAGGCC GAGACTGTGA TGGCAATGGA TATCCTGATC TGATTTGGGG GTCCCTTTGGT 1440  
 GTGACAAAG CTGTGGTATA CAGGGGCGGC CCCATGTTGT CCGCTAGTGC CTCCCTCACC 1500  
 ATCTTCCCGG CCATGTTCAA CCGCAGGAGG CGGAGCTGCA GCTTAGAGGG GAACCTGTG 1560  
 GCCTGCATCA ACCTTAGCTT CTGCCTCAAT GCTTCTGGAA AACACGTTGC TGACTCCATT 1620  
 GGTTCACAG TGAACCTCA GCTGGAAGTG CAGAAGCAGA AGGAGGGGGT ACGGCGGGCA 1680  
 CTGTTCTCTG CCTCCAGGCA GGCACCCCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT 1740  
 CGAGAGGATT GCAGAGAGAT GAAGATCTAC CTCAGGAACG AGTCAGAAAT TCGAGACAAA 1800  
 CTCTGCGCGA TTCACATCGC TCTCAACTTC TCCTTGGACC CCAAGCCCC AGTGGACAGC 1860  
 CACGGCTCA GCGCAGCCCT ACATTATCAG AGCAAGAGCC GGATAGAGGA CAAGGCTCAG 1920  
 ATCTTGTCTG ACTGTGGAGA AGACAACATC TGTGTGCTGT ACCTGCAGCT GGAAGTGT 1980  
 GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCC TGAACCTCAC TTTCCATGCC 2040  
 CAGAATGTGG GTGAGGGTGG CGCCTATGAG GCTGAGCTTC GGGTCACGCG CCCTCCAGAG 2100  
 GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACTTCT CCAGCCTGAG CTGTGACTAC 2160  
 TTTGCGGTGA ACCAGAGCGG CCGTCTGGTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA 2220  
 GCGAGTCTGT GGGGTGGCCT TCGGTTTACA GTCCCTCATC TCCGGGACAC TAAGAAAACC 2280  
 ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTGCAAGG CGACGTGGTT 2340  
 TCCTTTCGGC TCTCCGTGGA GGCTCAGGCC CAGGTCAACC TGAACGGTGT CTCACAGCCT 2400  
 GAGGCAATGC TATTCCAGT AAGGAGCTGG CATCCCGGAG ACCAGCTCA GAAGGAGGAG 2460  
 GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT 2520  
 AGCCAGGGTG TGCTGGAATC CAGCTGTCCC CAGGCTCTGG AAGGTACAGC GCTCCTATAT 2580  
 GTGACAGAG TTAGGGGACT CAATGACACC ACCAATCACC CCATTAAACC AAAGGGGCTG 2640  
 GAGTGGATC CCGAGGGTTC CCTGCACCA CAGCAAAAAC GGGAGAGTCC AAGCCGAGC 2700  
 TCTGCTTCTT CGGACCTCT GATCCTGAAA TGCCCGGAGG CTGAGTGT 2760  
 TGTGAGCTCG GGCCTCTGCA CCAACAAGAG AGCCAAAGTC TGCAATTGCA TTTCCGAGTC 2820  
 TGGGCCAAGA CTTTCTTGA GCGGAGCAC CAGCCATTTA GCCTGCAGTG TGAGGCTGTG 2880  
 TACAAGGCC TGAAGATGCC CTACCGAATC CTGCTCGGC AGCTGCCCA AAAAGAGCGT 2940  
 CAGGTGGCCA CAGCTGTGCA ATGGAACCAAG CGAGAAGGCA GCTATGGCGT CCACTGTGG 3000  
 ATCATATCC TAGCCTCTG GTTGGGCTC CTGCTCTCAT GTCTACTCAT CTACATCCTC 3060  
 TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCATGGA AAAAGCTCAG 3120  
 CTCAGGCTC CAGCCACTC TGATGCTGA



Seq ID NO: 387 Protein sequence  
Protein Accession #: NP\_002196.1

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
MGSRTPEPL HAVQLRWGPR RRPPLLPLLL LLLPPPPRVG GPNLDAEAPA VLSGPPGSPF 60
GFSVEFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCF WNASPTQCTP IEFDSKGSRL 120
LESSLSSESG BEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTICYLST 180
10     DNFTIRILEYA PCRSDFSWAA GQGYCQGGFS AEFTKTGRVV LGGPGSYFWQ GQILSATQSQ 240
      IAESYYPEYL INLVGQQLQT RQASSIYDDS YLGYSVAVGE PSGDDTDFPV AGVPKGNLTY 300
      GYVTIANGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DBLLVGAPLL MDRTPDGRPQ 360
      EVGRVYVYLQ HPAGIESPTPT LTLTGHDSEF RFGSSLTPLG DLDQDGYNDV AIGAPFGSET 420
      QQGVVVFVPG GPGGLGSKPS QVLQPLMAAS HTPDFPGSAL RGRGRDLGNG YPDLLVGSFG 480
15     VDKAVVYRGR PIVSASASLT IFPAMPNPEE RSCSLEGNPV ACINLSFCLN ASGKHVADSI 540
      GPTVELQLDW QKQKGGVRRR LPLASRQATL TQTLILQNGA REDCREMKIY LRNESEFRDK 600
      LSPHIALNPF SLDPQAPVDS HGLRPLHYQ SKSRIEDKQA ILLDCGEDNI CVPDLQLEVP 660
      GEQNHVYLGQ KNALMLTPHA QNVGEGGAYE AELRVTAPEE AEYSQVLRHP GNPSSLSCDY 720
      FAVNQSRLLV CDLGNPKMAG ASLWGLRPT VPHLRDTKKT IQPDFQILSK LNNSQSDVV 780
20     SFRLSVEAQA QVTNLGWSKP EAVLFPVSDW HPRDQPKKEE DLGPAVHEVY ELINQGPSII 840
      SQGVLELSCP QNLGQQQLLY VTRVTGLNCT TNHPINPKGL ELDPGSLHH QQKREAPSR 900
      SASGFPQLK CPEAEFRRLR CELGFLHQQE SQSLQLHFRV WAKTFLQREH QPFSLQCEAV 960
      YKALRMPYRI LPRQLPKER QVATAVQWTK AEGSYGVPLW IILAILPGL LLLGLLIYIL 1020
      YKLGFFKRSL PYGTAMEKAQ LKPPATSDA

```

Seq ID NO: 388 DNA sequence  
Nucleic Acid Accession #: NM\_002425  
Coding sequence: 26..1453

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
AAAGAAGSTA AGGGCAGTGA GAATGATGCA TCTTGCAATC CTGTGCTGTG TGTGTCGCC 60
AGTCTGCTCT GCGTATCTCT TGAGTGGGGC AGCAAAAGAG GAGGACTCCA ACAAGGATCT 120
35     TGCCACAGCA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAC AGTTTAGAAG 180
      AAAGGACAGT AATCTCATTT TTAATAAAAT CCAAGGAATG CAGAAGTCC TTGGGTGGA 240
      GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGGAGT 300
      TCTGAGCTTT GGTCACTTCA GCTCCTTTCC TGGCATGCGG AAGTGGAGGA AAACCCACCT 360
      TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTG ATTCTGCCAT 420
      TGAGAAAGCT CTGAAGTCT GGAAGAGGT GACTCCACTC ACATTCTCCA GGCTGTATGA 480
40     AGGAGAGGCT GATATAATGA TCTCTTTCG AGTTAAAGAA CATGGAGACT TTTACTCTTT 540
      TGATGGCCCA GACGACAGTT TGGCTCATGC CTACCCACCT GACCTGGGCG TTTATGGAGA 600
      TATTCACATT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATCTCT 660
      CGTTGCTGCT CATGAACCTG GCCACTCCCT GGGGCTCTTT CACTCAGCCA ACCTGAAGC 720
      TTTGATGTAC CCCTCTTACA ACTCATTCAC AGAGCTCGCC CAGTTCGCC TTTGCAAGA 780
45     TGATGTGAAT GGCATTCAGT CTCTCTACGG ACCTCCCCCT GCCTCTACTG AGGAACCCCT 840
      GGTGCGCACA AAATCTGTTC CTTCGGATC TGAGATGCCA GCCAAGTGTG ATCCTGCTTT 900
      GTCTCTCGAT GCCATCAGCA CTCTGAGGGG AGAATATCTG TTCTTTAAAG ACAGATATTT 960
      TTGGGGAAGA TCCCACTGCA ACCCTGAACC TGAATTTCAT TTGATTTCTG CATTTTGGCC 1020
      CTCTCTTCCA TCATATTGCG ATGCTGCATA TGAAGTTAAG AGCAGGGACA CCGTTTTTAT 1080
50     TTTTAAAGGA AATGAGTTCT GGGCCATCAG AGGAAATGAG GTACAAGCAG GTTATCCAAG 1140
      AGGCATCCAT ACCCTGGGTT TTCTCCAAAC CATAAGGAAA ATTGATGCAG CTGTTCTGA 1200
      CAGGAGAAAG AAGAAACACT ACTCTTTGCG AGCGGACAAA TACTGGAGAT TTGATGAAA 1260
      TAGCCAGTCC ATGGAGCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320
      GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTCCT TACTTCTTCA GTGGATCATC 1380
55     ACAGTTTGAG TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTAAACAGCTG 1440
      GTTACATTGC TAGGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTAATA 1500
      ATTATTATC TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCCTGCATG TCTGTGACT 1560
      GAAGAAGATG AGCCTTGCGA ATATCTGCAT GTGTCTAGAA GAATGTTTCT GGAATCTTCT 1620
      ACTTGCTTTT GAATTGCACT GAACAGAAAT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
60     ATGTATTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGCCTGTTC 1740
      CTT

```

Seq ID NO: 389 Protein sequence  
Protein Accession #: NP\_002416

```

65     1      11      21      31      41      51
      |      |      |      |      |      |
MHLAFLVLLC LPVCSAYPLS GAAKEEDSNK DLAQQYLEKY YNLEKDVQKQ RRKDSNLIVK 60
KIQGMQKFLG LEVTGKLDTD TLEVVRKPRC GVPDVGHFSS FPGMPKWRKT HLTIRIVNYT 120
70     PDLPRDAVDS ALEKALKVWE EVTPLTFSRL YEGEADIMIS FAVKEHGDYF SFDGPGHSLA 180
      HAYFPGPGLY GDHIFDDDEK WTEDASGTNL FLVAAHELGH SLGLFHSANT EALMYPLVNS 240
      FTSLAQFRLS QDDVNGIQSL YGPPFPASTZ PLVPTKSVPS GSEMPAKCDF ALSFDAISTL 300
      RGEYLFPRDR YFWRRSWHPN EPEFHLISAF WPSLPSYLD AYEVSNSRDV FIPKGNFPA 360
75     IRGNEVQAGY PRGIHTLGFPT IIRKIDAAV SDKEKKKTYF PAADKYWRFD ENSQSMQGF 420
      PRLIADDFPG VEPKVDVAVL AFGPFYFFSG SSQFEFDPNA RMVTHILKSN SWLHC

```

Seq ID NO: 390 DNA sequence  
Nucleic Acid Accession #: NM\_002421.2  
Coding sequence: 1..1409

```

80     1      11      21      31      41      51
      |      |      |      |      |      |
ATGCACAGCT TTCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTG ACACAGCTTC 60
CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAA 120
85     TACTACACC TGAAGAAATG TGGGAGGCAA GTTGAAAGC GGAGAAATAG TGGCCCACTG 180
      GTTGAAAAAT TGAAGCAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT 240
      GCTGAAACCC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCTGTATGT GGCTCAGTTT 300

```

GTCCTCACTG AGGGGAAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360  
 TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCAACTC 420  
 TGGAGTAATG TCACACCTCT GACATTACCC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480  
 ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540  
 CTGCTCATG CTTTTCAACC AGGCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600  
 GAAAGGTGGA CCAACAATT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGAAGTC 660  
 GGCCATTCTC TTGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC 720  
 ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT 780  
 GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840  
 AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900  
 TTCTACATGC GCACAAATCC CTCTACCCG GAAGTTGAGC TCAATTTTCT TTCTGTTTTT 960  
 TGCCCAACAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020  
 CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080  
 CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140  
 CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTG CTAACAAATA CTGGAGGTAT 1200  
 GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCTT 1260  
 GGAATTGGCC ACAAGTTTGA TGCAGTTTTT ATGAAAGATG GATTTTCTTA TTTCTTTTCT 1320  
 GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380  
 AATAGCTGGT TCAACTGCAG GAAAAATTAG

Seq ID NO: 391 Protein sequence  
 Protein Accession #: NP\_002412.1

1 11 21 31 41 51  
 MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60  
 VEKLKQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYYRIEN 120  
 YTPDLPRADV DHAIEKAPQL WSNVPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180  
 LAHAFQPGPG IGGDAHFDEB ERWTNNFREY NLRHVAHAEL GHSGLSHST DIGALMYPSTY 240  
 TFSGDVQLAQ DDIDIGIAIY GRSQNPVQPI GPQTPKACDS KLTFFDAITTI RGEVMMFFKDR 300  
 FYMRTNPFYP EVELNFIISV WPQLPNGLA AYEFADRDEV RFFKGNKYWA VQGNVLHGTY 360  
 PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFWANKYWRV DEYKRSMDPG YPKMLAHDFP 420  
 GIGHKVDVAV MKDGGFFYPFH GTRQYKDFPK TKRILTQKA NSWFNCRKN

Seq ID NO: 392 DNA sequence  
 Nucleic Acid Accession #: NM\_002421.2  
 Coding sequence: 1..1409

1 11 21 31 41 51  
 ATGCACAGCT TTCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGGTGTG ACACAGCTTC 60  
 CCAGCGACTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120  
 TACTACAACC TGAAGAATGA TGGGAGGCCAA GTTGAAGAGC GGAGAAATAG TGGCCAGTGT 180  
 GTTGAAGAAAT TGAAGCAAT GCAGGAATTC TTTGGGCTGA AAGTGACTGG GAAACCAGAT 240  
 GCTGAACACC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCTGTATGT GGCTCAGTTT 300  
 GTCTCACTG AGGGGAAACCC TCGCTGGGAG CAAACACATC TGACCTACAG GATTGAAAAAT 360  
 TACACGCCAG ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCCAACTC 420  
 TGGAGTAATG TCACACCTCT GACATTACCC AAGGTCTCTG AGGGTCAAGC AGACATCATG 480  
 ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540  
 CTGCTCATG CTTTTCAACC AGGCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600  
 GAAAGGTGGA CCAACAATT CAGAGAGTAC AACTTACATC GTGTTGCGGC TCATGCCCTC 660  
 GGCCATTCTC TTGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCCTAGCTAC 720  
 ACCTTCAGTG GTGATGTTCA GCTAGCTCAG GATGACATTG ATGGCATCCA AGCCATATAT 780  
 GGACGTTCCC AAAATCCTGT CCAGCCCATC GGCCCAAAA CCCCAAAAGC ATGTGACAGT 840  
 AAGCTAACCT TTGATGCTAT AACTACGATT CGGGGAGAAG TGATGTTCTT TAAAGACAGA 900  
 TTCTACATGC GCACAAATCC CTCTACCCG GAAGTTGAGC TCAATTTTCT TTCTGTTTTT 960  
 TGCCCAACAC TGCCAAATGG GCTTGAAGCT GCTTACGAAT TTGCCGACAG AGATGAAGTC 1020  
 CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080  
 CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140  
 CTTTCTGAGG AAAACACTGG AAAAACCTAC TTCTTTGTG CTAACAAATA CTGGAGGTAT 1200  
 GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCTT 1260  
 GGAATTGGCC ACAAGTTTGA TGCAGTTTTT ATGAAAGATG GATTTTCTTA TTTCTTTTCT 1320  
 GGAACAAGAC AATACAAATT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380  
 AATAGCTGGT TCAACTGCAG GAAAAATTAG

Seq ID NO: 393 Protein sequence  
 Protein Accession #: NP\_002412.1

1 11 21 31 41 51  
 MHSFPPLLLL LFWGVVSHSF PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60  
 VEKLKQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRWE QTHLTYYRIEN 120  
 YTPDLPRADV DHAIEKAPQL WSNVPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180  
 LAHAFQPGPG IGGDAHFDEB ERWTNNFREY NLRHVAHAEL GHSGLSHST DIGALMYPSTY 240  
 TFSGDVQLAQ DDIDIGIAIY GRSQNPVQPI GPQTPKACDS KLTFFDAITTI RGEVMMFFKDR 300  
 FYMRTNPFYP EVELNFIISV WPQLPNGLA AYEFADRDEV RFFKGNKYWA VQGNVLHGTY 360  
 PKDIYSSPGF PRTVKHIDAA LSEENTGKTY FFWANKYWRV DEYKRSMDPG YPKMLAHDFP 420  
 GIGHKVDVAV MKDGGFFYPFH GTRQYKDFPK TKRILTQKA NSWFNCRKN

Seq ID NO: 394 DNA sequence  
 Nucleic Acid Accession #: NM\_014331.2  
 Coding sequence: 1..1506

1 11 21 31 41 51

ATGGTCAGAA AGCCTGTTGT GTCCACCATC TCCAAAGGAG GTTACCTGCA GGGAAATGTT 60  
 AACGGGAGGC TGCCCTTCCCT GGGCAACAAG GAGCCACCTG GGCAGGAGAA AGTGAGCTG 120  
 AAGAGGAAAG TCACCTTACT GAGGGGAGTC TCCATTATCA TTGGCACCAT CATTGGAGCA 180  
 5 GGAATCTTCA TCCTCTCTAA GGGCGTGCTC CAGAACAAGG GCAGCGTGGG CATGCTCTG 240  
 ACCATCTGGA CGGTGTGTGG GGTCCGTGCA CTATTGTGAG CTTTGTCTTA TGCTGAATTG 300  
 GGAACAACCTA TAAAGAAATC TGGAGGTGAT TACACATATA TTTTGGAAAT CTTGGTCCA 360  
 TTACCAGCTT TTGTACGAGT CTGGGTGGAA CTCCTCATAA TACGCCCTGC AGCTACTGCT 420  
 GTGATATCCC TGGCATTGAG AGCTACATT CTGGAACCAT TTTTATTICA ATGTGAAATC 480  
 10 CCTGAACCTG CGATCAAGCT CATTACAGCT GTGGGCATAA CTGTAGTGAT GGTCTCTAAT 540  
 AGCATGAGTG TCAGCTGGAG CGCCCGATC CAGATTCTCT TAACCTTTTG CAAGCTCACA 600  
 GCAATTTCTG TAATTATAGT CCCTGGAGTT ATGCAGCTAA TTAAGGTCA AACGCAGAAC 660  
 TTTAAAGACG GGTTTTCAGG AAGAGATTCA AGTATTACGC GGTGGCCACT GGTCTTTTAT 720  
 TATGGAATGT ATGCATATGC TGGCTGGTTT TACCTCAACT TTGTTACTGA AGAAGTAGAA 780  
 15 AACCCCTGAA AAACCAATCC CTTTGAATA TGTATATCCA TGGCCATTGT CACCATTGGC 840  
 TATGTGCTGA CAAATGTGGC CTACTTTAGC ACCATTAACT CTGAGGAGCT GCTGCTTTCA 900  
 AATGCACTGG CAGTGACCTT TCTGAGCGG CTACTGGGAA ATTTCTCATT AGCAGTTCCG 960  
 ATCTTTGTTG CCTCTCCCTG CTTTGGCTCC ATGAACGGTG GTGTGTTGCG TGCTCCAGG 1020  
 TTATTCTATG TTGGCTCTCG AGAGGGTCAC CTTCCAGAAA TCCTCTCCAT GATTCTATGC 1080  
 CGCAAGCACA CCCTCTACC AGCTGTTATT CTTTGGCACC CTTTGACAAAT GATAATGCTC 1140  
 20 TTCTCTGAGG ACCTCGACAG TCTTTTGAAT TTCTCTCAGT TTGCCAGGTG GCTTTTATT 1200  
 GGGCTGGCAG TGCTGGCTGT GATTTATCTT CGATACAAAT GCCCAGATAT GCATCGTCTT 1260  
 TTCAAGGTGC CACTGTTTCT CCCAGCTTTG TTTTCTTCA CATGCTCTCT CATGGTTGCC 1320  
 CTTTCCCTCT ATTCGAGCCC ATTTAGTACA GGGATTGGCT TCGTCTATC TCTGACTGGA 1380  
 25 GTCCCTGGGT ATTATCTCTT TATTATATGG GACAAGAAAC CCAGTGGTGT TAGAATAATG 1440  
 TCAGAGAAAA TAAACAGAAC ATTACAAATA ATACTGGAAG TTGTACCAGA AGAAGATAAG 1500  
 TTATGAACCTA ATGGAATCTG GATCTTGGCA ATCTGCCCAA GGGGAGACAC AAAATAGGGA 1560  
 TTTTACTTTC ATTTCTGAA AGTCTAGAGA ATTACAACCT TGGTGATAAA CAAAGGAGT 1620  
 CAGTTATTTT TATTCATATA TTTTAGCATA TFCGAACCTA TTTCTAAGAA ATTTAGTTAT 1680  
 30 AACTCTATGT AGTTATAGAA AGTGAATATG CAGTTATCTT ATGAGTCCGA CAATCTCTGA 1740  
 GCTCTCTGATA CCTACCTATT GGGGTTAGGA GAAAAGACTA GACAATTACT ATGTGTCAT 1800  
 TCTCTACAAAC ATATGTTAGC ACGGCAAGA ACCTTCAAAAT TGAAGACTGA GATTTTCTG 1860  
 TATATATGGG TTTGTAAAG ATGTTTAC ACATACAGA TGTCTATACT GTGAAAAGTG 1920  
 TTTTCAATTG TGAATAAAG CATACATCAT GATTATGGCA AAGAGGAGAG AAGAAATTT 1980  
 35 ATTTTACATT GACATTGCAT TGCTTCCCTT TAGATACCAA TTTAGATAAC AAACACTCAT 2040  
 GCTTTAATGG ATTATACCCA GAGCATTGTT AACAAAGGTC AGTGGGGATT GTTGAATACA 2100  
 TTAAGAAAGA TGTCTAGGG GCTACTGTTT ATGAGACACA TCCAGGAGTT ATGTTTAAAT 2160  
 AAAAATCCTT GAGAAATTTT TATGTGAGAT GTTTTTCAT TCATTATCAG GAAGTTTATG 2220  
 TTATCTGCTA TTTTCTTTT TCACATCAGT TTGATCAGGA AAGTGATATA CACATCTTAG 2280  
 40 AGCAAGAGTT AGTTTGGTAT TAAATCCTCA TTAGAACAAC CACCTGTTTC ACTAATAACT 2340  
 TACCCCTGAT GAGTCTATCT AAACATATGC ATTTTAAGCC TTCAAATTAC ATTATCAACA 2400  
 TGAGAGAAAT AAACCAACAA GAAGATGTTT AAAATAATAG TCCCATATCT GTAATCATAT 2460  
 TCACATGCAA TGTTAGTAAT TCTGAAGTTT TTTAAATTTA TGGCTATTTT TACACGATGA 2520  
 TGAATTTTGA CAGTTTGTGC ATTTTCTTTA TACATTTTAT ATTCTTCTGT TAAATATCT 2580  
 45 CTTCAGATGA AACTGTCCAG ATTAATTAGG AAAAGGCATA TATTACATA AAAATTGCAA 2640  
 AAGAAATGTC GCTGTAAATA AGATTACAA CTGATGTTTC TAGAAAATTT CCACTTCTAT 2700  
 ATCTAGGCTT TGTCAATTAAT TTCCACACT TAATTATCAT TCAACTTGCA AAAGAGACAA 2760  
 CTGATAAGAA GAAATTTGAA ATGAGAACTT GTGGATAAGT GTTTGTGTTT AGAAGATGTT 2820  
 GTTTTGGCAG TATTAGAAA TACTGTGAGC CGGGCATGGT GGCTTACATC TGTAAATCCCA 2880  
 50 GCACCTTGGG AGGCTGAGGG GGTGGATCAC CTGAGGTGGG GAGTTCTAGA CCAGCCTGAC 2940  
 CAACATGGAG AAACCCCATC TCTACTAAAA ATACAAAATT AGCTGGGCAT GGTGGCAGAT 3000  
 GCTGTGAATC TCAGCTATTG AGGAGGCTGA GGCAGAGGAA TTGCTTGAAC CCGGAGGCG 3060  
 GAGGTTCAG TGAGCCAAAG TTGCACCACT GTACTCCAGC CTGGGTGACA AAGTCAGACT 3120  
 CCATCTCCAA AAAAAAAAAA AAAA

Seq ID NO: 395 Protein sequence  
 Protein Accession #: NP\_055146.1

1 11 21 31 41 51  
 MVRKPVVSTI SKGGVILQNV NGRPLSLGNK EPPGQEKVQL KRKVTLLRGV SIIIGTIIGA 60  
 GIFISPKGVL QNTGVSVMGL TIWTVCGVLS LFGALSYAEL GTTIKSGGH YTYILEVFGP 120  
 LPAPVRVWE LLIIRPAATA VISLAFGRYI LEFFFIQCEI PELAIKLITA VGITVVMVLN 180  
 SMSVWSARI QIFLTFCKLT AILIIIVPGV MQLIKGQTQN PKDAPSGRDS SITRLPLAPY 240  
 65 YGMYAYAGWP YLNFVTEVEE NPEKTIPLAI CISMATIGV YVLTVNAVYFT TINAEELLLS 300  
 NAVAVTFSEI LLGNFSLAVP IFVALSCFGS MNGGVFAVSR LPYVASREGH LPEILSMIHV 360  
 RKHTPLPAVI VLHPLTMIML FSGDLDSLILN FLSFARMLFI GLAVAGLIYL RYKCPDNHRP 420  
 FKVPLFIPAL PSFTCLPMVA LSLYSDPFST GIGFVITLTG VPAYLFLIIN DKPRWFRIM 480  
 SEKIIRTLQI ILEVVPEDK L

Seq ID NO: 396 DNA sequence  
 Nucleic Acid Accession #: NM\_006528  
 Coding sequence: 57..764

1 11 21 31 41 51  
 GCCGCCAGCG GCTTTCTCGG ACGCCTTGCC CAGCGGGCCG CCCGACCCCC TGCACCATGG 60  
 ACCCGCTCG CCGCTCGGG CTGTGATTG TGCTGCTTTT CCTGACGGAG GCTGCACTGG 120  
 80 GCGATGCTGC TCAGAGAGCA ACAGGAAATA ACGCGGAGAT CTGTCTCTCT CCCCTAGACT 180  
 ACGGACCCCTG CCGGGCCCTA CTTCTCGGT ACTACTACGA CAGGTACACG CAGAGCTGCC 240  
 GCCAGTCTCT GTACGGGGGC TCGAGGGGCA ACGCCAACAA TTTCTACACC TGGGAGGCTT 300  
 GCGAOGATGC TTGCTGGAGG ATAGAAAAAG TTCCCAAAGT TTGCGGCTG CAGTGTAGTG 360  
 TGGAGACCA GTGTGAGGG TCCACAGAAA AGTATTCTT TAATCTAAGT TCCATGACAT 420  
 85 GTGAAAAAT CTTTTCGGGT GGGGTGTCAC GGAACCCGAT TGAGAACAGG TTTCCAGATG 480  
 AAGCTACTTG TATGGGCTTC TGCGCACCAA AGAAAAATCC ATCATTTTGC TACAGTCCAA 540  
 AAGATGAGGG ACTGTGCTCT GCCAATGTGA CTGCTATTA TTTAATCCA AGATACAGAA 600  
 CCTGTGATGC TTTACCTAT ACTGGCTGTG GAGGGAATGA CAATAACTTT GTTAGCAGGG 660

AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGAAAAAGAA AAAGAAGATG CCAAGGCTTC 720  
 GCTTTGCCAG TAGAATCCGG AAAATTCGGA AGAAGCAATT TTAACATTC TTAATATGTC 780  
 ATCTTGTGTTG TCATTATGGC TTATTTGCCT TTATGGTTGT ATCTGAAGAA TAATATGACA 840  
 GCATGAGGAA ACAAATCATT GGTGATTAT TCACCACTTT TTATTAATAC AAGTCACTTT 900  
 TTTCAAAATTT TGGATTTTTT TATATATAAC TAGTGTCTAT TCAATGTGA GTCTACCAATT 960  
 TTTAATTTAT GGTTCACACTG TTTGTGAGAC GAATTCCTGC AATGCATAAG ATATAAAGC 1020  
 AAATATGACT CACTCATTTT TTGGGGTGTG ATTCTGTATT TCAGAAGAGG ATCATAACTG 1080  
 AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAAGGACTAG 1140  
 CC

Seq ID NO: 397 Protein sequence  
 Protein Accession #: NP\_006519

1 11 21 31 41 51  
 MDPARPLGLS ILLFLTEAA LGDAAQEPTG NNAEICLLPL DYGPCRALLL RYYIDRYTQS 60  
 CRQFLYGGCE GNANFYTWE ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFNLSSM 120  
 TCEKFFSGGC HRNRIENRFP DEATCMGFCA PKKIPSPCY S PKDEGLCSAN VTRYFNPRI 180  
 RTCDAPTYTG CGGNDNNFVS REDCKRACAK ALKKKKMPK LRFASRIKI RKKQF

Seq ID NO: 398 DNA sequence  
 Nucleic Acid Accession #: NM\_001508.1  
 Coding sequence: 1..1361

1 11 21 31 41 51  
 ATGGCTTCAC CCAGCCTCCC GGGCAGTGAC TGCTCCCAA TCATTGATCA CAGTCATGTC 60  
 CCCGAGTTTG AGGTGGCCAC CTGGATCAAA ATCAACCTTA TTCTGGTGA CCTGATCATC 120  
 TTCTGATAGG GCCTTCGGG GAACAGCGTC ACCATTCGGG TCACCCAGGT GCTGCAGAAG 180  
 AAAGGATACT TGCASAAGGA GGTGACAGAC CACATGGTGA GTTTGGCTTG CTGGACATC 240  
 TTGGTGTTC TCATCGGCAT GCCCATGGAG TTCTACAGCA TCATCTGGAA TCCCTGACC 300  
 ACGTCCAGCT ACACCTGTG CTGCAAGCTG CACACTTTC TTCTCGAGGC CTGCAGCTAC 360  
 GCTACGCTGC TGACCTGTG GAGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTCAACCC 420  
 TTCAAGTACA AGGCTGTGT GGGACCTTG CAGTGAAGC TGCTGATTGG CTTCGTCTGG 480  
 35 GTCACTCTCG CCTGTGTGGC ACTGCCCTG CTGTTTGCCA TGGTACTGA GTACCCCTG 540  
 GTGAACGTGC CAGCCACCG GGGTCTCACT TGCAACCGCT CCAGCACCCG CCACCAAGAG 600  
 CAGCCCGAGA CCTCAATAT GTCCATCTGT ACCAACCTCT CCAGCCTG GACCGTGTTC 660  
 CAGTCCAGCA TCTTCGGCGC CTTCTGGTC TACCTCTGG TCCTGCTCTC CTGAGCCTTC 720  
 ATGTGTGGA ACATGATGA GGTGCTCATG AAAAGCCAGA AGGGCTGCT GGCCTGGGGC 780  
 40 ACGCGCCTC GCAGCTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840  
 ACCATCATCT TCTGAGGCT GATTGTTGT ACATTGGCCG TATGTGGAT GCCCAACCAG 900  
 ATTCGAGGA TCATGCTGC GGCCAAACCC AAGCAGACT GGACGAGGTC CTACTTCCGG 960  
 GCGTACATGA TCTCTCTCC CTTCTCGAG ACGTTTTCT ACCTCAGCTC GGTCAATCAAC 1020  
 CCGCTCCTGT ACACGGTGT CTGCGAGCAG TTTCCGGCGG TGTTCTGTGA GGTGCTGTGC 1080  
 45 TGCCGCTGT GCTCTCAGCA CGCCAAACAC GAGAAGCGCC TGCGCTGACA TGCGCACTCC 1140  
 ACCACCGACA CGCCCGCTT TGTGCAGCG CCGTTGCTCT TCGCTCCCG CGCCAGTCC 1200  
 TCTGCAAGGA GAATGAGAA GATTTCCTTA AGCACTTTTC AGAGCGAGGC CGAGCCCGAG 1260  
 TCTAAGTCCC AGTCATTGAG TCTGAGTCA CTAGAGCCCA ACTCAGGCGC GAAACCGAGC 1320  
 AATTCTGCTG CAGAGAATGG TTTTCAGGAG CATGAAGTTT GA

Seq ID NO: 399 Protein sequence  
 Protein Accession #: NP\_001499.1

1 11 21 31 41 51  
 MASPSLPGSD CSQIIDHSHV PEFEVATWIK ITLILVYLII FVMGLLNSV TIRVTQVLQK 60  
 KGYLQKEVTD HMVSLACSDI LVFLIGPMPE FYSIIMNPLT TSSYTLSCKL HTFLPEACSY 120  
 ATLLHLVLTLS FERYIAICHP FRYKAVSGPC QVLLIGFVW VTSALVALPL LFAMSTEYPL 180  
 60 VNVPSHRGLT CNRSTRHHE QPETSNMISC TNLSSRWTF QSSIFGAFV YLVVLLSVAF 240  
 MCWMMQVLM KSQKSLAGG TRPQLRKSE SEESRTARRQ TIIPLRLIV TLAVCMMPNQ 300  
 IRRIMAAAKP KHDWTRSYFR AYMLLPFSE TFFYLSSVIN PLLYTVSSQ FRRVFVQVLC 360  
 CRLSLQHANH EKRLRVHAHS TDSARFVQR PLLFASRRQS SARTEKIFL STFQSEAEQ 420  
 SKSQSLSLES LEPNSGAKPA NSAAENGQFQ HEV

Seq ID NO: 400 DNA sequence  
 Nucleic Acid Accession #: NM\_006475.1  
 Coding sequence: 28..2538

1 11 21 31 41 51  
 AACGAACTG CAACGGAGAG ACTCAAGATG ATTCCCTTT TACCCATGTT TTCTCTACTA 60  
 TTGCTGCTTA TTGTTAACCC TATAAACGCC AACAATCATT ATGACAAGAT CTTGGCTCAT 120  
 AGTCGTATCA GGGGTGCGGA CCAAGGCCCA AATGCTGTG CCCTTCAACA GATTTTGGGC 180  
 75 ACCAAAAGA AATACTTCAG CACTTGTAG AACTGGTATA AAAAGTCCAT CTGTGGACAG 240  
 AAAACGACTG TTTTATATGA ATGTTGCCCT GGTATATGA GAATGGAAGG AATGAAGGC 300  
 TGCCCGACAG TTTTGGCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC 360  
 ACAAACGAGC GCTATCTGA CGCCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGATCC 420  
 TTCATCTACT TTGACCGAG TAATGAGGCT TGGACAACCT TGGATTCTGA TATCCGTAGA 480  
 GGTTTGGAGA GCAACGTGAA TGTGGAATTA CTGAATGCTT TACATAGTCA CATGATTAAT 540  
 80 AAGAGAATGT TGACCAAGGA CTTAAAAAAT GGCAATGATTA TTCCTTCAAT GTATAACAAT 600  
 TTGGGGCTTT TCATTAACCA TTATCCTAAT GGGGTGTGCA CTGTTAATTG TGCTCGAATC 660  
 ATCCATGGGA ACCAGATTGC AACAAATGGT GTTGTCCATG TCATTGACCG TGTGCTTACA 720  
 CAAATTTGGT CCTCAATTCA AGACTTCATT GAAGCAGAAG ATGACCTTTC ATCTTTTAGA 780  
 85 GCAGCTGCCA TCACATCGGA CATATTGGAG GCCCTTGGAA GAGACGGTCA CTTCAACTC 840  
 TTTGCTCCCA CCAATGAGGC TTTTGAGAAA CTTCCACGAG GTGCTCTAGA AAGGTTCTGT 900  
 GGAGACAAAG TGGCTTCCGA AGCTCTTATG AAGTACCACA TCTTAAATAC TCTCCAGTGT 960  
 TCTGAGCTTA TTATGGGAGG AGCAGTCTTT GAGACGCTGG AAGGAATATC AATTGAGATA 1020

GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGAACAA AAAGGATATT 1080  
 GTGACAAATA ATGGTGTGAT CCATTGATT GATCAGGTCC TAATTCCTGA TTCTGCCAAA 1140  
 CAGTGTATTG AGCTGGCTGG AAAACAGCAA ACCACCTTCA CGGATCTTGT GGCCCAATTA 1200  
 GGCTTGGCAT CTGCTCTGAG GCCAGATGGA GAATACACTT TGCTGGCACC TGTGAATAAT 1260  
 GCATTTCTCG ATGATACTCT CAGCATGGTT CAGCGCTCC TTAATAATAT TCTGCAGAA 1320  
 ACATATTATG AAGTAAAGT TGGCCTTAAT GAGCTTTACA ACGGGCAAAT ACTGGAAACC 1380  
 ATCGGAGGCA AACAGCTCAG AGTCTTCGTA TATCGTACAG CTGTCTGCAT TGAAAAATTA 1440  
 TGCAATGAGA AAGGGAGTAA GCAAGGGAGA AACGGTGGCA TTCACATAAT CCGCGAGATC 1500  
 ATCAAGCCAG CAGAGAAATC CCTCCATGAA AAGTTAAAC AAGATAAGCG CTTTAGCACC 1560  
 TTCTCAGGCC TACTTGAAGC TGCAGACTTG AAAGAGCTCC TGACACAACC TGGAGACTGG 1620  
 ACATTAATTG TGCCAACCAA TGATGCTTTT AAGGGAATGA CTAGTGAAGA AAAAGAAATT 1680  
 CTGATACGGG ACAAAAATGC TCTTCAAAAC ATCATTCTTT ATCACCCTGAC ACCAGGAGTT 1740  
 TTCATTGGAA AAGGATTGTA ACCTGGTGTG ACTAACATTT TAAAGACCAC ACAAGGAAGC 1800  
 AAAATCTTTC TGAAGAAGT AAATGATACA CTTCTGGTGA ATGAATTGAA ATCAAAAGAA 1860  
 TCTGACATCA TGACAAACAA TGGTGTAAAT CATGTGTAG ATAAACTCCT CTATCCAGCA 1920  
 GACACACCTG TTGGAATGTA TCAACTGCTG GAAATACTTA ATAAATTAAT CAAATACATC 1980  
 CAAATTAAGT TGTTCCTGGT TAGCACCTTC AAAGAAATCC CCGTGACTGT CTATCAAACT 2040  
 AAAATTATAA CCAAGTGTG GGAACCAAAA ATTAAGTGA TTGAAGGCAG TCTTCAGCCT 2100  
 ATTATCAAAA CTGAAGGACC CACACTAACA AAAGTCAAAA TTGAAGGTGA ACCTGAATTC 2160  
 AGACTGTATT AAGAGGTGA AACAACTAAT GAAGTGATCC ATGGAGAGCC AATTATTAAA 2220  
 AAATCACCCA AAATCATTGA TGGAGTGCCT GTGGAATAAA CTGAAAAAGA GACACGAGAA 2280  
 GAACGAATCA TTACAGGTCC TGAATAAAAA TACACTAGGA TTTCTACTGG AGGTGGAGAA 2340  
 ACAGAAGAAA CTCTGAAGAA ATTTGTTACA GAAGAGGTCA CCAAGGTCAC CAAATTCATT 2400  
 GAAGTGGTGA ATGGTCATTT ATTTGAAGAT GAAGAAATTA AAAGACTGCT TCAGGGAGAC 2460  
 ACACCGGTGA GGAAGTGTGA AGCCCAACAA AAAGTTCAGG GTTCTAGAAG ACGATTAAAG 2520  
 GAAGGTGCTT TCTAGTGAAG ATCCAAAAAC CAGAAAAAAA TGTTTATACA ACCCTAAGTC 2580  
 AATAACCTGA CCTTAGAAAA TTGTGAGAGC CAAGTTGACT TCAGGAACCT AAACATCAGC 2640  
 ACAAGAAGC AATCATCAAA TAATCTGAAA CACAAATTTA ATATTTTCTT TTCTGAATGA 2700  
 GAAACATGAG GGAATTTGTG GAGTTAGCCT CCTGTGGTAA AGGAATTGAA GAAATATATA 2760  
 CACCTTACAC CCTTTTTCAT CTTGACATTA AAAGTCTCGG CTAACCTTGG AATCCATTAG 2820  
 AGAAAAATCT TTGTGACCGA ATTCATTACA ATTCAAATCG AAGAGTTGTG AACTGTTATC 2880  
 CCATTGAAAA GACCGAGCCT TGTATGTATG TTATGGATAC ATAAATGCA CGCAAGCCAT 2940  
 TATCTCTCCA TGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTATA 3000  
 TCAAAAGGCT TTGCACATTT CTATATGAGT GGGTTTACTG GTAAATTATG TTATTTTATA 3060  
 CAACTAATTT TGTACTCTCA GAATGTTTGT CATATGCTTC TTGCAATGCA TATTTTATA 3120  
 TCTCAAAAGT TTCAATAAAA CCATTTTCCA GATATAAAGA GAATTACTTC AAATTGAGTA 3180  
 ATTCAGAAA ACTCAAGATT TAAGTTAAAA AGTGGTTTGG ACTTGGGAA

Seq ID NO: 401 Protein sequence  
 Protein Accession #: NP\_006466.1

1 11 21 31 41 51  
 MIPFLPMFSL LLLLVNPNIN ANNHYDKILA HSRIRGRDQG PNVCALQQLL GTKKKYFSTC 60  
 KNWYKKSICG QKTTVLYECC PGYMRMEGMK GCPAVLPIDH VYGLTGIVGA TTTQRYSDAS 120  
 KLREIEISGKG SFTYFAPSNE AMDNLDSDIR RGLSESNVVE LNLALHSEMI NKRLMTKDLK 180  
 NGMILPSMYN NLGLFINHYP NGVVTVNCAR IHHGNQIATN GVVHVIVRVL TQIGTSIQDF 240  
 IEAEDDLSSP RAAATISDIL BALGRDGEFT LFAPTNEAFE KLPBGVLERF MGDKVASEAL 300  
 MKYHILNTLQ CSSEIMGGAV FETLEGNITIE IGCDDGSITV NGIKMNVKRD IVTNNGVIHL 360  
 IQDVLIPDSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLSM 420  
 VQRLLKLILQ NEILKVRVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKSGKQG 480  
 RNGAIHIFRE IIKPAEKSLH EKLQDKRPS TFLSLLEAAD LKELLTPQGD WTLFPVTNDA 540  
 FRGMTSEKEE ILIRDKNALQ NIILYHLTPG VPIGKGFEPG VINILKTQOG SKILFLKEVND 600  
 TLLVNELEKSK ESDIMTNGV IHVVDKLLYP ADTPVGNLQL LEILNKLIKY IQIKFVRGST 660  
 FKEIPVTVYT TKIITKVVEP KIKVIBGSLQ PIKTEGPTL TKVKIEGEPE FRLIKBEGETI 720  
 TEVIBGEPIL KKYTKIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780  
 QEEVTKVTKP IEGGDGHLFE DEEIKRLQGG DTPVRKLQAN KKVQGSRRRL REGRSQ

Seq ID NO: 402 DNA sequence  
 Nucleic Acid Accession #: NM\_002416  
 Coding sequence: 40..417

1 11 21 31 41 51  
 ATCCAATACA GGAGTGACTT GGAACCTCAT TCTATCACTA TGAAGAAAAG TGGTGTCTCT 60  
 TTCTCTTGG GCATCATCTT GCTGGTCTG ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120  
 AAGGGTGCCT GTTCTCGCAT CAGCACCAAC CAAGGGAGCTA TCCACCTACA ATCTTGAAA 180  
 GACCTTAAAC AATTGCCCC AAGCCCTTCC TCGGAGAAAA TTGAAATCAT TGCTACACTG 240  
 AAGAAATGAG TTCAAACATG TCTAAACCCA GATTTCAGCAG ATGTGAAGGA ACTGATTAAA 300  
 AAGTGGGAGA AACAGGTGAG CCAAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAAA 360  
 AAGAAAGTTC TGAAGGTTCG AAAATCTCAA CGTTCTCGTC AAAAGAAGAC TACATAAGAG 420  
 ACCACTTCAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACCGCTATCA 480  
 TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAAT TTGACTAGAA AATTAAAAAC 540  
 ATTACTCTGA AATTGTAAC TAAAGTTAGAA AGTTGATTTT AAGAAATCAA ACGTTAAGAA 600  
 TTGTTAAAGG CTATGATTGT CTTTGTCTCT CTACCAACCA CCAGTTGAAT TTCATCATGC 660  
 TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCACCCAA CCACATCCCA 720  
 CTCACAACAG CTGCTCGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCCAGAGAG 780  
 TATCTGAGCG ACATGTGAGC AAGTCTTAAG CTTGTTAGCA TGCTGGTGAAG CCAAGCAGTT 840  
 TGAATTTGAG CTGACCTTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC 900  
 CTACAGGCTT CACACACAAT GTGCTGAGA GATTCTAGCT GATTGTTATT GGGTATCACC 960  
 ACTGGAGATC ACCAGTGTGT GGCCTTCAGA GCCTCCCTTC TGGCTTTGGA AGCCATGTGA 1020  
 TTCCATCTTG CCGCTCAGG CTGACCACTT TATTTCTTTT TGTTCCTCTT TGTCTCATTC 1080  
 AAGTCAGCTC TTCTCATCTC TACCACAATG CAGTGCCTTT CTTCTCTCCA GTGCACCTGT 1140  
 CATATGCTCT GATTATCTG AGTCAACTCC TTTCTCATCT TGTCCCCAAC ACCCCACAGA 1200  
 AGTGTCTTCT TCTCCCAATT CATCCTCACT CAGTCCAGCT TAGTTCAAGT CCGCTCTCTT 1260  
 AAATAAACCT TTTTGGACAC ACAAATTATC TTAATACTCC TGTTCCTACT GGTTCAGTAC 1320  
 CACATGGGTG AACACTCAAT GGTAACTATA TTCTTGGGTG TTTATCCTAT CTCTCCAAAC 1380

AGATTGTGAG CTCCTTGAGG GCAAGAGCCA CAGTATATTT CCCTGTTTCT TCCACAGTGC 1440  
 CTAATAATAC TGTGGAACATA GGTTTTAATA ATTTTITTAAT TGATGTTGTT ATGGGCAGGA 1500  
 TGGCAACACG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC TCCATGTTGG 1560  
 CTAGCCTCTG GTAACCTCTT ACTTATTATC TTCAGGACAC TCACTACAGG GACCAGGGAT 1620  
 GATGCAACAT CCTGTCTTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680  
 GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAATATA CAGTTTACCG 1740  
 AAAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800  
 CCAACCATAA AAAAATTCCT TTTCCCGAAG GAAAAGGGCT TTCTCAATAA GCCTCAGCTT 1860  
 TCTAAGATCT AACCAAGTAG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAATATG 1920  
 AGTTTTATTG TCCGTTTACT TGTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980  
 TCTCCCATGA AGAAAGGGAA CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040  
 TAGTGAAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAACCC TCCTTCCAGG 2100  
 GGAGGTTTCA TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTTAA CCTATACTCA 2160  
 CTTTCCCAAA TTGAATCACT GCTCACACTG CTGATGATT AGAGTGCTGT CCGGTGGAGA 2220  
 TCCCAACCGA ACCTCTTATC TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280  
 AAAAAATCAA GTGTTTCTA TAATTGAGAG TCTGTGACCC ACTTACCTTG CATCTCAG 2340  
 GTAGACAGTA TATAACTAAC AACCAAGAC TACATATTGT CACTGACACA CAGTTATAA 2400  
 TCATTATCA TATATATACA TACATGCATA CACTCTCAA GCAATAATT TTTCACTCA 2460  
 AACACGATT GACTGTGATA CCTTGTAAAT TGAATAATTT TCTTTGTAA AATAGAATGG 2520  
 TATCAATAAA TAGACCATTA ATCAG

Seq ID NO: 403 Protein sequence  
 Protein Accession #: NP\_002407

1 11 21 31 41 51  
 MKKSGVLFLG GIILLVLIGV QGTFVVRKGR CSCISTNOGT IHLQSLKDLK QPAPSPSCEK 60  
 IEIIATLKNQ VQTCLMPDSA DVKELIKWE KQVSQKKQK NGKKHQKKV LKVRKSQRSR 120  
 QKKT

Seq ID NO: 404 DNA sequence  
 Nucleic Acid Accession #: NM\_006670  
 Coding sequence: 85..1347

1 11 21 31 41 51  
 CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGGCGC GTCCAGGCCC 60  
 AGCTCCGGGG AAAAGCGAGC CGGATGCTT GGGGGGTGCT CCGGGGGCCC CGCGCGCGG 120  
 GACGGGCGTC TGGGCTGGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCTCGTCT 180  
 TCTCCCACTT CCTCGGCATC CTCTTCTTCC TCCTCGGGCG CGTTCCTGGC TTCCGCGGTG 240  
 TCGGCCAGCG CCGCGCTGCC GGACCAAGTG CCGCGCTGT GCGAGTGCTC CGAGGCGAGG 300  
 CGCAGAGTCA AGTGCGTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC 360  
 GTGCGCAACC TCTTCTTCTT CCGCAACCAG CTGGCGGTGC TCCTTGGCGG CGCTTTCGCC 420  
 CGCGGCGCGC CGCTGGCGGA GCTGGCGCGC CTCACCTCA GCGCGAGCGC CTTGGACGAG 480  
 GTGCGCGCGG GCGCCTTCCA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540  
 CCACTGGCGG ACCTCAGTCC CTTCGCTTTC TCGGCGACGA ATGCCAGCGT CTCGGCCCCC 600  
 AGTCCCTCTG TGAACCTGAT CCTGAACCAAC ATCGTCCCCC CTGAAGATGA GCGGCGAAGC 660  
 CGGAGCTTGG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCGGTGCACT GCAGGGGCTC 720  
 CGCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGATGT GCTGGCCCAA 780  
 CTGCCAGGCC TCAGGCACCT GGACTTAAGT AATAATTGCG TGGTGAGCCT GACCTACGTC 840  
 TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCACCTGG AGGACAATGC CCTCAAGGTC 900  
 CTTCAACATG GCACCTTGGC TGAGTTGCAA GGCTACACCC ACATTAGGGT TTTCTTGGAC 960  
 AACCAATCCCT GGTCTGCGTA CTGCCACATG GCAGACATGG TGACTGGCTT CAAGGAACAA 1020  
 GAGGTAGTGC AGGGCAAGA CCGGCTCACC TGTGCATATC CGGAAAAAT GAGGAATCGG 1080  
 GTCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140  
 CAAACCTCTT ATGTCTTCTT GGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCCTG 1200  
 GTTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAAACATCA AGATGCTGTC 1260  
 AGGGATCACA TGAAGGGTAC TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAAAC 1320  
 AACCTCAGTT CTAACCTGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380  
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440  
 TAGATACAAC GCACTTACG TAAAGCAGT GAAGGGGATT TGCTTCTTGT TTATGTAAG 1500  
 TTTCTCGGTG TGTCTGTATA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCTTC 1560  
 TTTCTTTTCT TGAACCTCT CAACACGTAT GGAGGGATT TTAGGTTTC AGCATGAACA 1620  
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680  
 ACAGATAGCA TTCAACAAAA GCTGCCTCAA CTTTTCGAG AAAAATACTT TATTCATAAA 1740  
 TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAAA AATTGCATCC TATAAAGTGC 1800  
 CTGCAGAGCT TACAGGCTC TTCAAAATAA CTCATGGTG CACAGGAGCA CCTGCATCCA 1860  
 AGAGCATGCT TACATTTTAC TGTCTGCAT ATTACAAAAA ATAACCTTGA ACTTCATAAC 1920  
 TTTCTTGACA AAGTAAATTA CTTTCTTGAT TGCAGTTTAT ATGAAAAATG ACTGATTTT 1980  
 TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAAATAAG 2040  
 ATTCTTAAAA GAA

Seq ID NO: 405 Protein sequence  
 Protein Accession #: NP\_006661

1 11 21 31 41 51  
 MPGGCSRGA AGDGRLLAR LALLVLLGWS SSSPTSSASS FSSSAPPLAS AVSAQPLPLD 60  
 QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL 120  
 AALNLSGSR LDEVRAAFEH LPSLRQLDLS HNPDLADLSP AFSGSNASVS APSPLVELIL 180  
 NHIVPPEDER QNRSEPMV V AALLAGRALQ GLRRLLEASN HFLYLPRLVL AQLPLSLRHL 240  
 LSNLSLVSLT YVSPFNLT L ESLHLEDNAL KVLHNGTLAE LQGLPHIRVP LNNPNWVDC 300  
 HMAHMTWLK ETEVVQKDR LTCAYPEKMR NRVLLELNSA DLDGDPILPP SLQTSYVFLG 360  
 IVLALIGAIF LVLVLYLNRKG IKKWMENIRD ACRDRMEGYH YRYRINADPR LTNLSNSDV

Seq ID NO: 406 DNA sequence  
 Nucleic Acid Accession #: Bos sequence

Coding sequence: 1..927

```

5 1 11 21 31 41 51
   | | | | |
   | | | | |
ATGCCTGGGG GGTGCTCCCG GGGCCCCGCC GCGGGGACG GCGGTCTGCG GCTGGCGCGA 60
CTAGGCTGGG TACTCCTGGG CTGGGTCTCC TCGTCTTCTC CCACCTCCTC GGCACTCCTC 120
TTCTCCTCCT CGGCGCCGCTT CCTGGCTTCC GCGGTGTCCG CCCAGCCCCC GCTGCGCGAC 180
CAGTGCCCGG CGCTGTGCGA GTGCTCCGAG CGAGCGCGCA CAGTCAAGTG CGTTAACCCG 240
AATCTGACCG AGGTGCCAC GACCTGCCC GCCTACGTGC GCAACCTCTT CCTTACCGGC 300
10 AACCAGCTGG CCAGCAACCA CTTCCTTTAC CTGCCGCGGG ATGTGCTGGC CCAACTGCCC 360
AGCCTCAGGC ACCTGGACTT AAGTAATAAT TCGCTGGTGA GCCTGACCTA CGTGTCTCTC 420
CGCAACCTGA CACATCTAGA AAGCCTCCAC CTGGAGGACA ATGCCCTCAA GGTCTCTCAC 480
AATGGCACCC TGGCTGAGTT GCAAGGTCTA CCCCACATTA GGGTTTTCTT GGACAACAT 540
CCTCGGTGCT CGACTCGCCA CATGGCAGAC ATGGTGACCT GGCTCAAGGA AACAGAGGTA 600
15 GTGCAAGGCA AAGACCGGCT CACCTGTGCA TATCCGAA AAATGAGGAA TCGGGTCTCT 660
TTGGAACCTA CAGTGTCTGA CCTGGACTGT GACCCGATTC TTCCCCATC CCTGCAAACT 720
TCTTATGTCT TCCTGGGTAT TGTTTTAGCC CTGATAGCG CTATTTCTCT CCTGGTTTGG 780
TATTTGAACC GCAAGGGGAT AAAAAAGTGG ATGCATAACA TCAGAGATGC CTGCAAGGAT 840
20 CACATGGAAG GGTATCATTA CAGATATGAA ATCAATGCGG ACCCCAGATT AACAAACCTC 900
AGTTCCTAAT CGGATGCTCT CGAGTGA

```

Seq ID NO: 407 Protein sequence  
Protein Accession #: Bos sequence

```

25 1 11 21 31 41 51
   | | | | |
   | | | | |
MPGGCSRGA AGDGRLLRL LALLLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQPPPLPD 60
QCPALCECSE AARTVKCVNR NLTEVFTDLP AYVRNLPLTG NQLASNHFLY LPRDVLALPL 120
30 SLRHLDSLNN SLVSLTYVSF RNLTHLESLE LEDNALKVLH NGTLAEQLQL PHIRVFLDNN 180
PWVDCDHMD MVTWLKTEV VQKDRLTCA YPKMRNRVL LELNSADLDC DPILPPSLQT 240
SYVPLGIPLA LIGAIFLLVL YLNRKGIKKW MENIRDACRD HMEGYHYRYE INADPRLLNL 300
SSNSDVLE

```

Seq ID NO: 408 DNA sequence  
Nucleic Acid Accession #: NM\_000095.1  
Coding sequence: 26..2299

```

40 1 11 21 31 41 51
   | | | | |
   | | | | |
CAGCACCCAG CTCCCGGCCA CGGCCATGGT CCCCACACCC GCCTGCGTTC TTCTGCTCAC 60
CTTGCTGCTC CTGCGCGCGT CCGGACAGGG CCAGAGCCCC TTGGGCTCAG ACCTGGGCCC 120
GCAGATGCTT CGGGAACATG AGGAAACCAA CGCGCGCGTG CAGGACGTGC GGGACTGGCT 180
CGGGCAGCAG GTCCAGGAGA TCACGTTCTT GAATAACACG GTGATGGAAT GTGACGCGTG 240
CGGGATGCAG CAGTCAGTAC GCACCGGCTT ACCCAGCGTG CGGCCCCGTC TCCACTGCGC 300
45 GCGCGGCTCT TGCTTCCCGG GGTGGGCTTG CATCCAGACG GAGAGCGCGG GCCGCTGCGG 360
GCCCTGCCCC CGGGGCTTCA CGGGCAACGG CTGCGACTGC ACCGACGTCA ACGAGTGCAA 420
CGCCACCCCC TGCTTCCCCG GAGTCCGCTG TATCAACACC AGCCCGGGGT TCCGCTGCGA 480
GGCTTGCCCG CGGGGTGACA GCGGCCCCAC CCACACGGGC GTGGGGCTGG CTTTCGCGAA 540
50 GGCACCAACG CAGGTTTGCA CGACATCAA CGAGTGTGAG ACCGGGCAAC ATAAGTCCGT 600
CCCCAATCTC GTGTGCATCA ACACCCGGGG CTCTTTCCAG TGCGGCCGCT GCCAGCCCCG 660
CTTCGTGGGG GACACAGCGT CCGGCTGCCA GCGCGCGCGA CAGCGCTTCT GCCCCGACGG 720
CTGCGCCAGC GAGTGCCACG AGCATGCAGA CTGCGTCTTA GAGCGCGATG GCTGCGGCTC 780
GTGCGTGTGT CGGCTTGGGT GGGCCGGCAA CGGGATCCTC TGTGTTGCGG ACACGTGACT 840
55 AGAOGGCTTC CGGACGAGA AGCTGCGCTG CCGGAGCGCG CAGTCCGCTA AGGACAACTG 900
CGTGACTGTG CCAACTCAG GGCAGGAGGA TGTGACGCG GATGGCATCG GAGACGCGTG 960
CGATCOGGAT GCCGACGGGG ACGGGGTCCC CAATGAAAAG GACAACTGCC CGCTGGTGGC 1020
GAACCCAGAC CAGCCCAACA CGGACGAGGA CAAGTGGGCG GATGCGTGGC ACAACTGCCG 1080
GTCCACAGAG AACGACGACC AAAAGGACAC AGACCAAGAC GGCCCGGGCG ATGCGTGCGA 1140
60 CGACGACATC GACGCGGACC GGATCGCAA CCAGGCGGAC AACTGCCCTA GGTACCCCAA 1200
CTCAGACCAG AAGGACAGTG ATGGCGATGG TATAGGGGAT GCCTGTGACA ACTGTCCCA 1260
GAAGAGCAAC CGGATCAGG CGGATGTGGA CCACGACTTT GTGGGAGATG CTTGTGACAG 1320
CGATCAAGAC CAGGATGGAG ACGGACATCA GGACTCTCGG GACAACTGTC CCAAGGTGCC 1380
TAACAGTGCC CAGGAGGACT CAGACCAAGA TGGCCAGGGT GATGCTTGCG ACGACGACGA 1440
65 CGACAATGAC GAGATCCCTG ACAGTCCGGA CAACTGCGCG CTGGTGCTTA ACCCGGCCCA 1500
GGAGGACGCG GACAGGGAAG GGTGGGCGA CGTGTGCCAG GACGACTTTG ATGCAGACAA 1560
GGTGGTAGAC AAGATCGACG TGTGTCCGGA GAACGCTGAA GTACGCTCA CCGACTTCAG 1620
GGCCTTCCAG ACAGTCTGTC TGGACCCGGA GGGTGACGCG CAGATTGACC CCAACTGGGT 1680
GGTGCTCAAC CAGGGAAGGG AGATCGTGCA GACAATGAAC AGGACCCGAC GCTTGGCTGT 1740
GGGTACACT GCCTTCAATG CGTGGGACTT CGAGGGCAGC TTCCATGTGA ACAAGTTCAC 1800
70 GGATGACGAC TATGCGGCTT TCATCTTTGG CTACCAAGAC AGCTCCAGCT TCTACGTGGT 1860
CATGTGGAAG CAGATGAGAC AAACGTATTG GCAGGCGAAC CCCTTCCGTG CTGTGGCCGA 1920
GCCTGGCATC CAACTCAGG CTGTGAAGTC TTCCACAGGC CCCGGGGAAC AGCTGCGGAA 1980
CGCTCTGTGG CATACAGGAG ACACAGAGTC CCAGGTGGCG CTGCTGTGGA AGGACCCGCG 2040
AAACGTGGGT TGGAGGAGCA AGAAGTCTTA TCGTTGGTTC CTGCAGCACC GCGCCCAAGT 2100
75 GGGCTACATC AGGGTGGGAT TCTATGAGGG CCCTGAGCTG GTGGCCGACA GCAACGTGGT 2160
CTTGACACCA ACCATGCGGG GTGGCCGCTT GGGGGTCTTC TGCTTCTCCC AGGAGAACAT 2220
CATCTGGGCC AACCTGGGTT ACCGCTGCAA TGACACCATC CCAGAGGACT ATGAGACCCA 2280
TCAGCTGGCG CAAGCTTAGG GACCAAGGAG AGGACCCGCG GGATGACGAGC CACCCCTACC 2340
80 GCGGCTGGAT GGGGCTCTG CACCCAGCCC AAGGGGTGGC CGTCTGAGG GGGAGGTGAG 2400
AAGGGCTCAG AGAGGACAAA ATAAAGTGTG TGTGCAAGG

```

Seq ID NO: 409 Protein sequence  
Protein Accession #: NP\_000086.1

```

85 1 11 21 31 41 51
   | | | | |
   | | | | |
MVFDTACVLL LTLAALGASG QGQSPLGSDL GPQMLRELQE TNAALQDVDR NLRQQVREIT 60

```

WO 02/086443

PCT/US02/12476

FLKNTVMEDD ACGMQQSVRT GLPSVRPLLE CARGFPCPGV ACIQTESGGR CGPCPAGFTG 120  
 NGSHCTDVNE CNAHPCFPRV RCINTSPGFR CEACPPGYSG PTHQGVGLAF AKANKQVCTD 180  
 INECETQGHN CVPNSVCINT RGSFQCGPCQ PGFVGDQASG CQRGAQRFCP DGSFSECHHE 240  
 ADCVLERDGS RSCVCRVWMA GNGILCGRDT DLDGFFDEKL RCPEPQCRKD NCVTVNSGQ 300  
 EDVDRDQIGD ACDDPADDGG VFNEKDNCPV VRNPDQRNTD EDKWDGACDN CRSQKNDQDK 360  
 DTDQDGRGDA CDDIDGDRJ RNQADNCPRV FNSDQKDSOG DGIQDADNDC PQKSNPDQAD 420  
 VDHDFVGDAC DSDQDQDGG HQDSRDNCPT VFNSAQEDSD HDQGDGACDD DDDNDGVFDS 480  
 RDNCRLVFNP QGEDADRGV GDVQDQDFDA DKVVDKIDVC PENAEVTLTD FRAFTQTVLD 540  
 PEGDAQIDPN WVVLNQGREI VQTMNSDPGL AVGYTAFNGV DFEGTFHVNT VTDDYAGFI 600  
 PGYQSSSFY VVMWQMEQT YWQANPFRV AEPGIQLKAV KSTGFGEBQL RNALWHTGDT 660  
 ESQVRLLEWD PRNVGWKDK SYRWFLQHRP QVGYIRVRFY EGPELVADSN VVLDITMRGG 720  
 RLGVFCFSQE NIIWANLRYR CNDTIPEDYE THQLRQA

Seq ID NO: 410 DNA sequence  
 Nucleic Acid Accession #: NM\_001565.1  
 Coding sequence: 67..363

1 11 21 31 41 51  
 | | | | | |  
 20 GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCCAGTCTC 60  
 AGCACCATGA ATCAAACTGC GATTCTGATT TGCTGCCTTA TCTTTCTGAC TCTAAGTGGC 120  
 ATTCAAGGAG TACCTCTCTC TAGAACCCTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180  
 CCTGTTAATC CAAGGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA 240  
 25 CGTGTGAGA TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300  
 TCGAAGGCCA TCAAGAATTT ACTGAAGACA GTTAGCAAGG AAATGTCTAA AAGATCTCCT 360  
 TAAAACCCAG GGGGAGCACA ATCGATGCAG TGCTTCCAAG GATGGACCAC ACAGAGGCTG 420  
 CCTCTCCCAT CACTTCCCTA CATGAGATAT ATGTCAAGCC ATAATTGTTC TTAGTTTGCA 480  
 GTTACACTAA AAGGTGACCA ATGATGGTCA CCAATCAGC TGCTACTACT CCGTAGGAA 540  
 30 GGTAAATGTT CATCATCTTA AGCTATTGAG TAATAACTCT ACCCTGGCAC TATAATGTAA 600  
 GCTCTACTGA TGCTCTATGT TCTTAGTGGA TGTCTGACC CTGCTTCAA TATTTCCTCTC 660  
 ACCTTTCCCA TCTTCCAAGG GTACTAAGGA ATCTTTCTGC TTTGGGGTTT ATCAGAATTC 720  
 TCAGAACTCT AATAACTTAA AAGGTATGCA ATCAAACTCT CTTTATAAAG AATGCTCTTT 780  
 ACTTCATGGA CTCTCACTGC CATCTCCCA AGGGGCCCAA ATTCTTTCAG TGGCTACCTA 840  
 35 CATACAAATC CAAACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAGTATT 900  
 CTTATTAAAT GAAAGACTGT ACAGAGTATA AGTCTTAGAT GTATATATTT CCTATATTGT 960  
 TTTCACTGTA CATGGAATTA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA 1020  
 TTTTAAAAAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG 1080  
 TTTTCAATA AAAATGAGT ACTCTCCTGG AATATTAAAG

Seq ID NO: 411 Protein sequence  
 Protein Accession #: NP\_001556.1

1 11 21 31 41 51  
 | | | | | |  
 45 MNQTALICC LIFLTLGSIQ GVPLSRTVRC TCISISNPQV NPSLEKLEI IPASQFCPRV 60  
 EIIATMKKKK EKRCINPESK AIKNLLKAVS KEMSKRSP

Seq ID NO: 412 DNA sequence  
 Nucleic Acid Accession #: XM\_057014  
 Coding sequence: 143..874

1 11 21 31 41 51  
 | | | | | |  
 55 GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCCTCGGAG 60  
 CGCGGGGAGG CAGAGCGCTG ACCAGGTTCC TCTCTCGGT CTCTCCGCC TCCAGCTCCG 120  
 CGCTGCCCGG CAGCGCGGAG CCATCGGACC CCAGGGCCCC GCGCGCTCCC CGCAGCGGCT 180  
 CGCGGGCCTC CTGCTGCTCC TGCTGCTGCA GCTGCCCGCG CGCTCGAGCG CCGCTGAGAT 240  
 CCCCAGGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300  
 60 AATGTGCTTA CAAGGGCCAG CAGGAGTGCC TGTGCGAGAC GGGAGCCCTG GGGCCAATGG 360  
 CATTCCGGGT ACACCTGGGA TCCAGGTGCG GGATGGATT AAAGGAGAAA AGGGGGAATG 420  
 TCTGAGGGA AGCTTTGAGG AGTCTGGGAC ACCCAACTAC AAGCAGTGT CATGGAGTTC 480  
 ATTGAATTAT GGCATAGATC TTGGGAAAAT TGCGGAGTGT ACATTACAA AGATGCGTTC 540  
 AATAGTGTCT CTAAGAGTTT TGTTCAGTGG CTCACCTCGG CTAAGATGCA GAAATGCATG 600  
 65 CTGTGAGCTT TGTGATTCCA CATTCAATGG AGCTGAATGT TCAGGACCTC TTCCCATGGA 660  
 AGCTATAATT TATTTGACCC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTTCATG 720  
 CACTTCTTCT GTGGAAGGAC TTGTGAAGG AATTGGTGCT GGATTAGTGG ATGTGTCTAT 780  
 CTGGGTTGGC ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840  
 TTCTCGCATC ATTATTGAAG AACTACCCAA ATAAATGCTT TAATTTTCAT TTGCTACCTC 900  
 70 TTTTATTATT ATGCTTGGGA ATGTTCACT TAAATGACAT TTTAATAAG TTTATGTATA 960  
 CATCTGAATG AAAAGCAAGG CTAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020  
 TTTAAATCTA GCATTATCCA TTTTGCTTCA ATCAAAAGTG GTTTCATAT TTTTITTAGT 1080  
 TGGTAGAAT ACTTCTTCCA TAGTCACATT CTCTCAACCT ATAATTGGGA ATATTGTGT 1140  
 GGTCTTTTGT TTTTCTCTT AGTATAGCAT TTTTAAAAAA ATATAAAAGC TACCAATCTT 1200  
 75 GTACAAATTT GTAAATGTTA AGAATTTT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260  
 CAACCTTAAA AAAAAA AAAA

Seq ID NO: 413 Protein sequence  
 Protein Accession #: XP\_057014

1 11 21 31 41 51  
 | | | | | |  
 80 MRPGGPAASP QRLRGLLLLL LLQLPAPSSA SEIPKGGKKA QLRQREVVDL YNGMCLQGPA 60  
 GVPGRDGSPP ANGIPTPGI PGRDGFPGKE GECLRESFEE SWTPNYKQCS WSSLNWYIDL 120  
 85 GKLAECTFTK MRSNSALRLV FSGSLRLKCR NACQQRWYFT FNGAECSSGL PIEAIIVLDQ 180  
 GSPFNSTIN IHRTSSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIEE 240  
 LPK



WO 02/086443

Seq ID NO: 414 DNA sequence  
Nucleic Acid Accession #: XM\_084007  
Coding sequence: 138..2405

PCT/US02/12476

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
CTCGTGCCGA ATTCCGGCAG AGACCGCGTG TTCGCGCCTG GTAGAGATTT CTCGAAGACA 60
CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCGG GGCGGTGGGA CAACGAGGCC 120
GCGGAGACGA AGGCCCAATG GCGAGGAAGT TATCTGTAAT CTGTATCTGT ACCTTTGGCC 180
10    TCTCTGTGTC AAATCCCTTT CATGAACATA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA 240
AAATTAGTCC GAATTTGGGA TCTGGCATTA ATGTTGACTT GGCAATTTCG ACACGGCAAT 300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAAATA TTTCTTGTCA GTTGAAGGGT 360
TCAGAAAAAT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAAATCCAT ATACACCATG 420
15    ACCACGACCA TCATCTCAGC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG 480
AGCATCACTC AGACCCAGAG CATCACTCTG ACCATGATCA TCACTCCCAC CATATCATG 540
CTGCTTCTGG TAAAAATAAG CGAAAGCTC TTTGCCGAGA CCATGACTCA GATAGTTTCA 600
GTAAGATATC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA 660
GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACACTG 720
20    TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAATCTTTCC 780
CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCAATC AAAGAGCCGG GTGAGCCGGC 840
TGGCTGGTAG GAAAAATAAT GAATCTGTGA GTGAGCCCGG AAAAGGCTTT ATGTATTCCA 900
GAAACACAAA TGAAAAATCT CAGGAGTGT TCAATGCATC AAGCTACTG ACATCTCATG 960
GCATGGGCAAT GCAATTCCTG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
25    TCACCAAAAT TGATGCTAGA TCTGTCTGTA TTCAATCAAG TGAAAAGAAG GCTGAAATCC 1080
CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA 1140
TCAGTTTCTT GTCTCTGCTG GGGGTTATCT TAGTGCTCTT CATGAATCGG GTGTTTTTCA 1200
AAATTTCTCT GAGTTTCTCT GTGGCAGTGG CGGTGGGAC TTTGAGTGGT GATGCTTTTT 1260
TACACCTTCT TCCACATCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAGAAACCAG 1320
30    CAATGGAAAT GAAAGAGAGA CCACTTTTCA GTCACTGTCT TCTCAAAAC ATAGAAGAAA 1380
GTGCTTATTT TGATCCACG TGGAAAGGTC TAACAGCTCT AGGAGCGCTG TATTTTATGT 1440
TTCTTGTTGA ACATGTCCTC ACATTGATCA AACAATTAA AGATAAGAAG AAAAAGAATC 1500
AGAAGAAACG TAAAAATGAT GATGATGTGG AGATTAGAA GCAGTTGTCC AAGTATGAAT 1560
35    CTCAACTTTC AACAATGAGT GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTATC 1620
GAGCAGACTC ACAGAGAGCC TCCACTTTG ATTCTCAGCA GCCTGCAATC TTGGAAGAAG 1680
AAGAGGTGAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
GGTGCAAGAA TAAATGCCAT TCACATTTCC ACATACACT CGGCCAGTCA GACGATCTCA 1800
40    TTCACCAACA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCAA AACCACCATC 1860
CTCAGACTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGG GTCCGCCATT 1920
TGGCTGTGAT GGTGATATG GGTGATGCCC TGCACAATTT CAGCGATGGC CTAGCAATTC 1980
GTGCTGCTTT TACTGAAGCG TTATCAAGTG GTTTAAGTAC TTTGTTGCT GTGTTCTGTC 2040
ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA 2100
AGCAGGCTGT CCTTTATAAT GCATGTCTAG CCATGCTGGC GTATCTTGA ATGGCAACAG 2160
45    GAATTTTCAT TGGTCATPAT GCTGAAATG TTTCTATGTA GATATTTGCA CTACTGCTG 2220
GCTTATTCAT GTATGTTGCT CTGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
GTGACCATGG ATGATAGCGC TGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
GTTTGTGAAT TATGTTACTT ATTTCATAT TTGAACATA AATCGTGTTC CGTATAAAT 2400
50    TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCAT 2460
AGGAGATGTA GTTTGATATG TGTACTATGC AGCGTTTAAA GTTAGTGGT TTTGTGATT 2520
TTGTATTGAA TATGCTGTCT TGTACAAAG TCAGTTAAAG GTACGTTTAA ATATTAAAT 2580
TATCTATCTT TGGATGATAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG 2640
55    TAAACAGAGG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAATAATGT CTTTAAATGCT 2700
TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACGTCTGG 2760
TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA 2820
AGCAAGAGAA TAAAGGAGAA AAGAGAGGAA TCTGAGAAIT GGGGAGGCAT AGATTCTTAT 2880
60    AAAAAACACA TAAATTTGTA TAAATTAGAG GGGAGAAAIT TAGAATTAAG TATAAAAGG 2940
CAGAAATAGT ATAGAGTACA TTCAATTAAC ATTTTGTICA GGATTAATTC CCGTAAAAAC 3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTAACT TTTGTATAAT ACAGAAATCT 3060
AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTTGAA TTTCAAATG 3120
70    TTCGTGGGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT 3180
TATGCGCAAG TTATATACCA CCAAAAGCTG TATGACTGGA TGTCTCGGTT ACCTGGTTTA 3240
80    CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAAACT ACACTAAGTA 3300
TCATTTGATT CGATTACAGA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTTG 3360
GAGCAATTGT CTTTATATAC GGTACTGTAG CCATCTAGG CCTGTCTGTG GCATCTCTTA 3420
65    GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G

```

Seq ID NO: 415 Protein sequence  
Protein Accession #: XP\_084007

```

70      1      11      21      31      41      51
      |      |      |      |      |      |
MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPFW ESGINVDLAI STRQYHLQQL 60
FYRYGENNSL SVBGFRLKQ NIGIDKIKRI HHHHDHDS DREHSDHER HSDHEHSDH 120
EHSDEHDS HHHHAAGSKN KKKALCPDHD SDSSGKDPFN SQKGGAHRPE HASGRNVKD 180
75    SVSASEVTST VYNTVSEGT FLETIETPRP GKLPKDVSS STPPSVTSKS RVSRLAGRKT 240
NESVSEPRKG FMYSRNTNEN PQECFNASKL LSHGMGIQV PLNATEPNYL CPAIINQIDA 300
RSLIHTSEK KAEIPPKTYS LQIAWVGFI AISIISFLSL LGVILVPLMN RVFPKPLPS 360
LVALAVGTLG GDAPLHLLPH SHASHHSHS HEAPAMEMKR GPLPSHLSQ NIEESAYFDS 420
TWKGLTALGG LYFMELVEHV LTLIKQPKDK KKKNQKPKEN DDDVEIKQL SKYESQLSTN 480
80    BEKVDTDRT EGYLRADSQE PSHFDSQQA VLEEEVMIA HAHPOEVYNE VPRGCKNKC 540
HSHFDHLLGQ SDDLHSHHD YHILHSHHD QNHHPHSHSQ RYSREELKDA GVATLAWMVI 600
MGDGLHNFSD GLAIGAAFTG GLSSGLSTSV AVFCHPELPE LGDPFVLLKA GNTVKQAVLY 660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS 720
RWGYFFLQNA GMLLPGFIML LISIFPKIV FRINP

```

Seq ID NO: 416 DNA sequence  
Nucleic Acid Accession #: NM\_015419.1  
Coding sequence: 1..8487

	1	11	21	31	41	51	
5	ATGCCCAAGC	GCGCGCACTG	GGGGGGCCCTC	TCCGTGGTGC	TGATCCTGCT	TTGGGGCCAT	60
	CCGCGAGTGG	CGCTGGCCTG	CCCCCATCCT	TGTGCTGCT	ACGTCCCCAG	CGAGGTCCAC	120
	TGCACGTTCC	GATCCCTGGC	TTCCGTGCCC	GCTGGCATTG	CTAGACACGT	GGAAAGAATC	180
	AATTGGGGT	TTAATAGCAT	ACAGGCCCTG	TCAGAAACCT	CATTGTCAGG	ACTGACCAAG	240
	TTGGAGCTAC	TTATGATTC	CGGCAATGAG	ATCCCAAGCA	TCCCGATGGS	AGCTTTAAGA	300
10	GACCTCAGCT	CTCTTCAGST	TTTCAAGTTC	AGCTACAACA	AGCTGAGAGT	GATCACAGGA	360
	CAGACCTCTC	AGGGTCTCTC	TAACTTAATG	AGGCTGCACA	TTGACCACAA	CAAGATCGAG	420
	TTTATCCACC	CTCAAGCTTT	CAACGGCTTA	ACGTCTCTGA	GGCTACTCCA	TTTGGAAAGG	480
	AATCTCCTCC	ACCAGCTGCA	CCCCAGCACC	TTCTCCACGT	TCACATTTTT	GGATTATTTT	540
	AGACTCTCCA	CCATAAGGCA	CCTCTACTTA	GCAGAGAACA	TGGTTAGAAC	TCTTCTGTCC	600
15	AGCATGCTTC	GGAAACATGCC	GCTTCTGGAG	AATCTTTACT	TGCAGGGAAA	TCCGTGGACC	660
	TGCGATTGTG	AGATGAGATG	GTITTTGGAA	TGGGATGCAA	AATCCAGAGG	AATTCTGAAG	720
	TGTAAAGAGG	ACAAAGCTTA	TGAAGGCGGT	CAGTTGTGTG	CAATGTGCTT	CAGTCCAAGG	780
	AAGTTGTACA	AACATGAGAT	ACACAAGCTG	AAGGACATGA	CTTGTCTGAA	GCCTTCAATA	840
20	GAGTCCCCTC	TGAGACAGAA	CAGGAGCAGG	AGTATTGAGG	AGGAGCAGAA	ACAGGAAGAG	900
	GATGTTGGCA	GCCAGCTCAT	CCTGGAGAAA	TTCCAACCTG	CCCACTGGAG	CATCTCTTTG	960
	AATATGACCG	ACGAGCACGG	GAACATGGTG	AACTTGGTCT	GTGACATCAA	GAACCAATG	1020
	GATGTTGTACA	AGATTCACTT	GAACCAAAAG	GATCCTCCAG	ATAATTGACAT	AAATGCAACA	1080
	GTTCCTTTGG	ACTTTGAGTG	TCCAATGACC	CGAGAAAAC	ATGAAAAGCT	ATGGAAATTG	1140
	ATAGCATACT	ACAGTGAAGT	TCCCGTGAAG	CTACACAGAG	AGCTCATGCT	CAGCAAGAGC	1200
25	CCCAGAGTCA	GCTACCACTG	CAGGCAAGAT	GCTGATGAGG	AAGCTCTTTA	CTACACAGGT	1260
	GTGAGAGCCC	AGATTCTTGC	AGAACCAGAA	TGGGTCTATG	AGCCATCCAT	AGATATCCAG	1320
	CTGAACCGAC	GTCAAGATAC	GGCCAAGAA	GTGCTACTTT	CCTACTACAC	CCAGTATTCT	1380
	CAAAACAATAT	CCACCAAGAA	TACAAGGCGG	GCTCGGGGCA	GAAGCTGGGT	AATGATTGAG	1440
	CTAGTTGGAG	CTGTGCAAG	AGATCAGACT	GTCTGTGAAG	GGGGTCCATG	CCAGTTGAGC	1500
30	TGCAACGTGA	AGGCTTCTGA	GAGTCCATCT	ATCTTCTGGG	TGCTTCCAGA	TGGCTCCATC	1560
	CTGAAGCGCC	CCATGGATGA	CCGACAGCAG	AAGTTCTCCA	TTCTCAGCAG	TGGCTGGCTG	1620
	AGGATCAAGT	CCATGGAGCC	ATCTGACTCA	GGCTTGTACC	AGTGCAATTG	TCAAGTGAGG	1680
	GATGAAATGG	ACCGCATGGT	ATATAGGGTA	CTTGTGCAAT	CTCCCTCCAC	TCAGCCAGCC	1740
	GAGAAAGACA	CAGTGACAA	TGGCAAGAAC	CCAGGGGAGT	CGGTGACATT	GCCTTGCAAT	1800
35	GCCTTTCGCA	CTGTGCAAG	CCACCTTAGC	TGGATTCTTC	CAACACAGAG	GATAATTAAT	1860
	GATTTGGCTA	ACACATCACA	TGTATACATG	TTGCCAAATG	GAACCTTTTC	CATCCCAAG	1920
	GTCCAACTCA	GTGATATGSG	TTACTACAGA	TGTGTGGCTG	TCAACACAGA	AGGGGCGAGC	1980
	CATTTTACGG	TGGGAATCAC	AGTGACCAAG	AAAGGCTCTG	GCTTGCCATC	CAAAAGAGGC	2040
40	AGACGCCGAG	GTGCAAGGCG	TCTTTCCAGA	GTCAAGAGAG	ACATCGTGGG	GGATGAAGGG	2100
	GGCTCGGGCA	TGGGAGATGA	AGAGAACACT	TCAAGGAGAC	TTCTGCATCC	AAAGGACCAA	2160
	GAGGTGTTCC	TCAAAACAAA	GGATGATGCC	ATCAATGGAG	ACAAGAAAGC	CAAGAAAGGG	2220
	AGAAGAAAGC	TGAACCTCTG	GAAGCATTCG	GAAGAAAGAA	CAGAGACCAA	TGTTGCAGAA	2280
	GGTGCGCAGAG	TGTTTGAATC	TAGACGAAGG	ATAAACATGG	CAAAACAACA	GATTATCCCG	2340
	GAGCGCTGGG	CTGATATTTT	AGCCAAAGTC	CGTGGGAAAA	ATCTCCCTAA	GGGACAGAA	2400
45	GTACCCCGGT	TGATTAATAA	CACAAGTCTC	CCATCTCTGA	GCCTAGAAGT	CACACCACCT	2460
	TTCTCTGCTG	TTTCTCCCCC	CTCAGCATCT	CCTGTGCAGA	CAGTAACCA	TGCTGAAGAA	2520
	TCTCTAGCAG	ATGTACTCTG	ACTTGGTGAA	GAAGAGCAGG	TTTTGGGTCA	CATTTCCTCA	2580
	GCCAGCATGG	GGCTAGAACA	CAACCACAAT	GGAGTTATTC	TTGTTGAACC	TGAAGTAACA	2640
	AGCACACCTC	TGGAGGAAGT	TGTTGATGAC	CTTTCTGAGA	AGACTGAGGA	GATAACTTCC	2700
50	ACTGAAGGAG	ACCTGAAGGG	GACAGCAGCC	CCTACACTTA	TATCTGAGCC	TTATGAACCA	2760
	TCTCTACTCT	TGCACACATT	AGACACAGTC	TATGAAAGAG	CCACCCATGA	AGAGACGGCA	2820
	ACAGAGGGTT	GGTCTGCAGC	AGATGTTGGA	TCGTCAACAG	AGCCCAATC	CAGTGAGTAT	2880
	GAGCCTCCAT	TGGATGCTGT	CTCCTTGGCT	GAGTCTGAGC	CCATGCAATA	CTTTGACCCA	2940
	GATTGAGAGA	CTAAGTCACA	ACCAGATGAG	GATAAGATGA	AAGAAGACAC	CTTTGCACAC	3000
55	CTTACTCCAA	CCCCCACCAT	CTGGGTTAAT	GACTCCAGTA	CATCAGAGTT	ATTGAGGAT	3060
	TCTACTATAG	GGGAACCCAG	TGTCCCAGGC	CAATCACATC	TACAGGAGCT	GACAGACAAC	3120
	ATCCACCTTG	TGAAAGATAG	TCTAAGCACT	CAAGACACCT	TACTGATTAA	AAAGGGATAG	3180
	AAAGAGATGT	CTCAGACACT	ACAGGGAGGA	AATATGCTAG	AGGGAGACCC	CACACACTCC	3240
	AGAAGTTCTG	AGAGTGAGGG	CCAAGAGAGC	AAATCCATCA	CTTTGCCCTG	CTCCACACTG	3300
60	GGTATAATGA	GCAGTATGTC	TCCAGTTAAG	AAGCCTGGCG	AAACCAACAG	TGGTACCTTC	3360
	CTAGACAAGG	ACACCAACA	AGTAACAACA	ACACCAAGGC	AAAAAGTTGC	TCCGTCTATC	3420
	ACCATGAGCA	CTCACCTTTC	TCGAAGGAGA	CCCAACGGGA	GAAGGAGATT	ACGCCCAAC	3480
	AAATTCGGCC	ACCGGCACAA	GCAACCCCA	CCCAACACTT	TTGCCCATC	AGAGACTTTT	3540
	TCTACTCAAC	CAACTCAAGC	ACCTGACATT	AAGATTTCAA	GTCAAGTGGG	GAGTTCTCTG	3600
65	GTTCTTACAG	CTTGGGTGGA	TAAACACAGT	AATACCCCA	AACAGTTGGA	AATGGAGAAG	3660
	AATGCGAAGC	CCACATCCAA	GGGAACACCA	CGGAGAAAAC	ACGGGAAGAG	GCCAAACAAA	3720
	CATCGATATA	CCCCCTTCTAC	AGTGAGCTCA	AGAGGCTCCG	GATCCAGGCC	CAGCCCTTCT	3780
	CCAGAAAATA	AACATAGAAA	CATTGTTACT	CCCAGTTGAG	AAACTATAGT	TTTGCTTAGA	3840
	ACTGTTTCTC	TGAAACCTGA	GGGCCCTTAT	GATTCTTTAG	ATTACATGAC	AAOCCACAGA	3900
70	AAAAATATAT	CATCTTACCC	TAAAGTCCAA	GAGACACTTC	CAGTCACATA	TAAACCCACA	3960
	TCAGATGGAA	AAGRAATTAA	GGATGATGTT	GCCACAATA	TTGACAAAAC	TAAAGTGAC	4020
	ATTTTAGTCA	CTGGTGAATC	AATTACTAAT	GCCATACCAA	CTTCTGCTC	CTTGGTCTCC	4080
	ACTATGGGAG	AATTTAAGGA	AGAACTCTCT	CCTGTAGGCT	TTCCAGGAAC	TCCAACCTGG	4140
	AATCCCTCAA	GGAAGGCCCT	GCCTGGGAGG	CTACAGACAG	ACATACTCTG	TACCACTTCT	4200
75	GGGAAAATC	TTACAGACCC	TCCCTTCTT	AAAGAGCTTG	AGGATGTGGA	TTTCACTTCC	4260
	GAGTTTTTGT	CCTCTTTGAC	AGTCTCCACA	CCATTTTACC	AGGAGAAGC	TGGTTCTTCC	4320
	ACAACCTCTC	CAAGCATAAA	AGTGGAGGTG	GCTTCAAGTC	AGGCAGAAAC	CACCAACCTT	4380
	GATCAAGATC	ATCTTGAAGC	CAGTGTGGCT	ATTCTCTTTT	CTGAAACTAG	ACCAAGAAAT	4440
	CACACCCCTA	CTGCTGCCCG	GATGAAGGAG	CCAGCATCTT	CGTCCCATAC	CACAAATCTC	4500
80	ATGCTTTTGG	GACAAACAC	CACCACTAAG	CCAGCACTTC	CCAGTCCAG	AATATCTCAA	4560
	GCATCTAGAG	ATTCCAAGGA	AAATGTTTTT	TTGAATTATG	TGGGGAATCC	AGAAACAGAA	4620
	GCAACCCGAG	TCAACAATGA	AGGAACACAG	CATATGTCTG	GGCCAAATGA	ATTATCAACA	4680
	CCCTCTTCCG	ACGGGATGTC	ATTTAACTTG	TCTACAAAGC	TGGAATTGGA	AAAGCAAGTA	4740
	TTTGGTAGTA	GGAGTCTACC	AOGTGGCCCA	GATAGCCCAAC	GCCAGGATGG	AAGAGTTCAT	4800
	GCTTCTCATC	AACTAACACG	AGTCCCTGCC	AAACCCATCC	TACCAACAGC	AACAGTGAGG	4860
85	CTACCTGAAA	TGTCCACACA	AAGCGCTTCC	AGATACCTTG	TAACTTCCCA	GTCACTCTGT	4920
	CAGTGGACCA	ACAAACCCGA	AATAACTACA	TATCCTTCTG	GGGCTTTGCC	AGAGAACAAA	4980
	CAGTTTACAA	CTCCAAGATT	ATCAAGTACA	ACAATTCCTC	TCCCATTTGCA	CATGTCCAAA	5040

	CCGAGCAITTC	CTAGTAAGTT	TACTGACCGA	AGAACTGACC	AATTCATATG	TTACTCCAAA	5100
	GTGTTTGGAA	ATAACACAT	CCCTGAGGCA	AGAAACCCAG	TTGGAAAGCC	TCCAGTCCA	5160
	AGAATTCTCT	ATTATTCCAA	TGGAAGACTC	CCTTTCTTTA	CCAACAAGAC	TCTTTCTTTT	5220
5	CCACAGTTGG	GAGTACCCCG	GAGACCCCG	ATACCCACTT	CTCCTGCCCC	AGTAATGAGA	5280
	GAGAGAAAAG	TTATTCCAGG	TTCTTACAAC	AGGATACATT	CCCATAGCAC	CTTCCATCTG	5340
	GACCTTGGCC	CTCGGACACC	TCCGTTGTTG	CACACTCCGC	AGACCAAGGG	ATCACCCCTCA	5400
	ACTAACTTAC	AGAATATCCC	TATGGTCTCT	TCCACCCAGA	GTCTATCTCT	CTTTATAACA	5460
	TCTTCTGTCC	AGTCCCTCAG	AAGCTTCCAC	CAGAGCAGCT	CAAAGTTCTT	TGCAGGAGGA	5520
10	CCTCCTGCAT	CCAAATTCCTG	GTCTCTTGGG	GAAAAGCCCG	AAATCCTCAC	CAAGTCCCCA	5580
	CAGACTGTGT	CCGTACACGC	TGAGACAGAC	ACTGTGTTC	CCTGTGAGGC	AACAGGAAAA	5640
	CCAAAGCCTT	TCGTTACTTG	GACAAAGGTT	TCCACAGGAG	CTCTATGAC	TCCGAATACC	5700
	AGGATACAAC	GGTTTGAAGT	TCTCAAGAAC	GGTACCTTAG	TGATACGGAA	GGTTCAGTA	5760
	CAAGATCGAG	GCCAGTATAT	GTGCACCGCC	AGCAACCTGC	ACGGCCTGGA	CAGGATGGTG	5820
	GTCTTGTCTT	CGGTACCGCT	GCAGCAACCT	CAAATCCTAG	CCTCCCACTA	CCAGGACGTC	5880
15	ACTGTCTACC	TGGGAGACAC	CATTGCAATG	GAGTGTCTGG	CCAAAGGGAG	CCGAGCCCCC	5940
	CAAAATTCCT	GCACTCTCCC	TGACAGGAGG	GTGTGGCAAA	CTGTGTCCCG	CGTGAGAGGC	6000
	CGCATCACCC	TGCACGAAAA	CCGAGCCCTT	TCCATCAAGG	AGGCGTCTCT	CTCAGACAGA	6060
	GGCGTCTATA	AGTGGCTGGC	CAGCAATGCA	GCAGGCGGGG	ACAGCTGGC	CATCCGCTGG	6120
20	CAGTGGGGG	CAGTGGGGGG	CGTTATCCAC	CAGGAGAAGC	TGGAGAACAT	CTCGCTGCCC	6180
	CGGGGGCTCA	GCATTACAT	TCACTGCAC	GCCAAAGGCTG	CGCCCTGCCC	CAGCGTGGCC	6240
	TGGGTGCTCG	GAGGAGGTAC	CCAGATCCGC	CCCTCGCAGT	TCTTCAAGGG	GAACCTGTGT	6300
	GTTCCTCCCA	ACGGGAGCCT	CTACATCCGC	AACCTCGCGC	CCAAGGACAG	CGGGCGCTAT	6360
	GAGTGGCTGG	CCGCAACCTT	GGTAGGCTCC	CGCGCAGGA	CGGTGCAGCT	GAACGTGCAG	6420
25	CGTGAGCAG	CCAAGCGCGG	CATCAGCGGC	ACCTCCCGCG	GGAGGACGGA	CGTCAGGTAC	6480
	GAGGGAACCC	TCAAGCTCGA	GTGACAGGCC	TGGGGGAGCC	CCTGGCCGCG	CATCCTCTGG	6540
	AGGCTGCCGT	CCAGAGGAGT	GATCGACGCG	CTCTTCAGTT	TTGATAGCAG	AATCAAGGTG	6600
	TTTGCCCAATG	GGAACCTGGT	GGTGAATCA	GTGACGGACA	AAGATGCCGG	AGATTACCTG	6660
	TGCGTAGCTC	GAAATTAAGT	TGGTGTATG	TACGTGGTGC	TCAAGTGGGA	TGTGGTGATG	6720
30	AAACCGGCCA	AGATTGAACA	CAAGGAGGAG	AACGACCACA	AAGTCTTCTA	CGGGGTGAC	6780
	CTGAAAGTGG	AGTGTGTGGC	CACCGGGCTT	CCCAATCCCG	AGATCTCTCG	GAGCCTCCCA	6840
	GACGGGAGTC	TGGTGAACAT	CTTCATGACG	TGCGATGACA	GCGGTGAGC	CACCAAGCGC	6900
	TATGTGGTCT	TCAACAAATG	GACACTCTAC	TTTAAACGAG	TGGGGATGAG	GGAGGAAGGA	6960
	GACTACACCT	GCTTTGTCTG	AAATCAGGTC	GGGAAGGACG	AGATGAGAGT	CAGAGTCAAG	7020
35	GTGGTGACAG	CGCCCGCCAC	CATCCGGAAC	AAGACTTACT	TGGCGGTCCA	GGTGCCCTAT	7080
	GGAGAGCTGG	TCACTGTAGC	CTGTGAGGCC	AAAGGAGAAC	CCATGCCCAA	GGTGACTTGG	7140
	TTGTCCCAAA	CCAAAGAGGT	GATCCCAACC	TCTCTGAGA	AGTATCAGAT	ATACCAAGAT	7200
	GGCAGTCTCC	TTATTCAAGG	AGCCAGCGGT	TCTGACAGCG	GCAACTACAC	CTGCTGGTCT	7260
	AGGAACAGCG	CGGGAGAGCG	TAGGAAGACG	GTGTGGATTG	ACGTCAACGT	CCAGCCACCC	7320
40	AAGATCAACG	GTAACCCCAA	CCCCATCACC	ACCGTGGCGG	AGATAGCAGC	CGGGGGCAGT	7380
	CGGAAACTGA	TTGACTGCAA	AGCTGAAGGC	ATCCCCACCC	CGAGGGGTGT	ATGGGCTTTT	7440
	CCGAGGGTGG	TGGTCTCTCC	AGCTCCATAC	TATGGAAACC	GGATCATCTG	CCATGGCAAC	7500
	GGTTCCTCTG	ACATCAGGAG	TTTGAGGAAG	AGCGACTCOG	TCCAGCTGGT	ATGCATGGCA	7560
	CGCAACGAGG	GAGGGGAGGC	GAGGTTGATC	GTGCAGCTCA	CTGTCTGAGA	GCCCATGGAG	7620
45	AAACCCATCT	TCCACGACCC	GATCAGCGAG	AAGATCAACG	CCATGGCGGG	CCACACCATC	7680
	AGCCTCAACT	GCTCTGCGCG	GGGGACCCCG	ACACCCAGCC	TGGTGTGGGT	CCTTCCCAAT	7740
	GGCACCGATT	TGCAGAGTGG	ACAGCAGCTG	CAGCGCTTCT	ACCACAAGGC	TGACGGCATG	7800
	CTACACATTA	GGGCTCTCTC	CTCGGTGGAC	GCTGGGGCCT	ACCGCTGGGT	GGCCCGCAAT	7860
	GGCGCTGGCC	ACACCGGAGG	GCTGGTCTCC	CTGAAGGTGG	GACTGAAGCG	AGAAAGCAAC	7920
50	AAGCAGTATC	ATAACCTGGT	CAGCATCATC	AATGGTGAGA	CCCTGAAGCT	CCCTTGCAAC	7980
	CCTCCCGGGG	CTGGGCGAGG	ACGTTTCTCC	TGACGCTCC	CCAATGGCAT	GCATCTGGAG	8040
	GGCCCCCAAA	CCCTGGGACG	CGTTTCTCTT	CTGGACAATG	GCACCCCTAC	GGTTCGTGAG	8100
	GCCTCGGTGT	TTGACAGGGG	TACCTATGTA	TGCAGGATGG	AGACGGAGTA	CGGCCCTTGG	8160
	GTCAACGAGA	TCCCGGTGAT	TGTGATCGCC	TATCCTCCCC	GGATCACCAG	CGAGCCCAAC	8220
55	CGGTCATCT	ACACCCGGCC	CGGGAACACC	GTGAACTGA	ACTGCATGGC	TATGGGGATT	8280
	CCCAAGCTG	ACATCACGTG	GGAGTTACCG	GATAAGTGGC	ATCTGAAGCG	AGGGGTTGAG	8340
	GCTCGTCTGT	ATGGAAACAG	ATTTCTTCAC	CCCCAGGAGT	CAGTACCAT	CCAGCATGCC	8400
	ACACAGAGAG	ATGCGGGCTT	CTACAAGTGC	ATGGCAAAAA	ACATTCCTGG	CAGTGACTCC	8460
	AAAAACAATT	ACATCCAGCT	CTTCTGAAAT	GTGGATTCCA	GAATGATTCG	TTAGGAAGCT	8520
60	ACAAACAAAG	GGGGTTTGTA	AGGGAAGCCA	GGTTGGGGAA	TAGGAGCTCT	TAAATATAGT	8580
	GTCAAGTGC	ATGGTGGGCT	CTGGTGGGTT	TCAAGTTGAG	GTGATCTGT	ATCTACAATT	8640
	GTGGGAAAA	GGAGCAATG	CAGACACGAG	AAGGAGGGCT	CAGCCTTGCT	GAGACACTTT	8700
	CTTTTGTGTT	TACATCATGC	CAGGGGCTTC	ATTCAAGGTG	TCTGTGCTCT	GACTGCAATT	8760
	TTTCTTCTTT	TGCAATATGCC	ACTCGACTGC	CTTCATAAGC	GTCCATAGGA	TATCTGAGGA	8820
65	ACATTTCATCA	AAAAATAGCC	ATAGACATGA	ACAAACACCTC	ACTACCCCAT	TGAAGACGCA	8880
	TCACCTAGTT	AACCTGCTGC	AGTTTTTACA	TGATAGACTT	TGTTCCAGAT	TGACAAGTCA	8940
	TCTTTCAGTT	ATTTCTCTG	TCACCTTCAA	ACTCCAGCTT	GCCCAATAAG	GATTTAGAAC	9000
	CAGAGTGACT	GATATATATA	TATATATTTT	AATTCAAGAT	TACATACATA	CAGCTACCAT	9060
	TTTATATGAA	AAAAGAAAAA	CATTTCTTCC	TGGAACCTAC	TTTTTATATA	ATGTTTTATA	9120
70	TATATATTTT	TTCTTTTCAA	ATCAGACGAT	GAGACTAGAA	GGAGAAATAC	TTTCTGTCTT	9180
	ATTAATAATTA	ATAAATATT	GGTCTTTACA	AGACTTGGAT	ACATTACAGC	AGACATGGAA	9240
	ATATAATTTT	AAAAAATTC	TCTCCAACCT	CCTTCAAAAT	CAGTCACCA	TGTTATATTA	9300
	CCTTCTCCAG	GAACCCCTCA	GTGGGGAAGG	CTGCGATATT	AGATTTCTCT	GTATGCAAGG	9360
	TTTTTGTGTA	AAGCTGTGCT	CAGAGGAGGT	GAGAGGAGAG	GAAGGAGAAA	ACTGCATCAT	9420
75	AACCTTACAG	AATGAATCT	AGAGTCTTCC	CCGAAAGGCC	CAGAAACTTC	TCTGCAGTAT	9480
	CTGGCTGTGC	CATCTGGTCT	AAGGTGGCTG	CTTCTTCCCC	AGCCATGAGT	CAGTTTGTGC	9540
	CCATGAATAA	TACACGACCT	GTATTTTCCA	TGACTGCTTT	ACTGTATTTT	TAAGTCAAT	9600
	ATACTGTACA	TTTGATAATA	AAATAATATT	CTCCCAAAAA	AAAAA		

Seq ID NO: 417 Protein sequence

Protein Accession #: NP\_056234.1

	1	11	21	31	41	51	
85	MPKRAHWGAL	SVVLLILWGH	PRVALACPHP	CACYVPSEVH	CTFRSLASVP	AGIARHVERI	60
	NLGFNSIQAL	SETSPAGLTK	LELLMIHNGE	IPSIDPGLR	DLSSLQVFKF	SYNKLRVITG	120
	QTLQGLSNLM	RHIDHNKIE	FIHPQAFNGL	TSRLRLHLEG	NLLHLQHPST	FSTPTFLDYP	180
	RLSTIRHLYL	AENMVRTLPA	SMLRNMPLE	NLYLQGNPWT	CDCEMRWPLE	WDAKSRGILK	240

CKKOKAYEGG QLCAMCFSPK KLYKHEIHLK KDMTCLKPSI ESPLRQNRSR SIEEBEQBEE 300  
 DGGSQLILEK FQLPQWSISL NMTDEHGMV NLVCDIKKPM DVYKIHNLQT DPPDIDINAT 360  
 VALDFECFMT RBNYEKLWKL IAYYSEVPVK LHRELMLSKD PRVSYQYRQD ADEEALYYTG 420  
 VRAQLAEPE WMQPSIDIQ LNRROSTAKK VLLSYTYQYS QTISTKQTRQ ARGRSWVMIE 480  
 PSGAVQRQDT VLEGGPCQLS CNVKASESPS IFWVLPDGS I LKAPMDDEPS KFSILSSGWL 540  
 RIKSMEPSDS GLYQCIQAVR DEMDRMVYRV LVQSPSTQPA EKDVTVTIGKN PGESVTLFCN 600  
 ALAIPEARLS WILPNRRRIIN DLANTSHVYM LPNGTSLIPK VQVSDSGYYR CVAVNQOGAD 660  
 HPTVGTITVK KSGGLPSKRG RRPKAKALSR VREDIVEDEG GSGMGDEBENT SRRLHLHPKDQ 720  
 EVELTKKDDA INGDKKAKKG RRKLKLWKHS EKEPETNVAE GRRVFESRRR INMANKQINP 780  
 ERWADILAKV RGNKLPKGTG VPPLIKTTSP PSLSLEVTPP FPAVSPSPAS PVQTVTSAEE 840  
 SSADVPLLGE EEHVLTGTTSS ASMGLEHNNH GVILVEPEVT STPLEEVVDD LSEKTEEITS 900  
 TEGDLKGTAA PTLISEPYEP SPITLTLDTV YEKPTEETHA TEGWSAADVG SSPEPTSSEY 960  
 EPPLDAVSLA ESEPMQYFDP DLETKSQPDE DRMKEDTFAH LTPTFTIWNV DSSTSQLFED 1020  
 STIGSEGVPG QSHLQGLTDN IHLVKSSLSL QDTLLIKKGM KEMSQTLQGG NMLEGDPHNS 1080  
 RSSESEGGES KSITLPDSTL GIMSSMSPVK KPAETTVGTL LDKDTTIVTT TPRQKVPASS 1140  
 TMTSTHPSRRR FNGRRRLRPN KFRHRHKQTP PTFAPSETP STQPTQAPDI KISSQVSSSL 1200  
 VPTAWVDNTV NTPKQLEMEK NAEPTSKGTP RRGHGKRPNK HRYTPSTVSS RASGSKPSPS 1260  
 PENKHNRTVT PSETILLER TVSLKTEGYP DSDLYMTTR KLYSSYPKVQ ETLPVTYKPT 1320  
 SDGKIKDDV ATNVKHKSD ILVTGESITN AIPTSRSLVS TMGEFKEESS PVGPGPTPTW 1380  
 NPSRTAQGR LQTDIPVITS GENLTDPLLL KELEDVDFTS EPLSSLTVST PFHQEAGSS 1440  
 TTSSISIKVEV ASSQAEITTL DQDLETTVA ILLSETRPON HTPTAARMKE PASSSPSTIL 1500  
 MSLGQTTTITK PALPSPRISQ ASRDSKENVF LNYVGNPETE ATPVNNEGTQ HMSGPNELST 1560  
 PSSDRDAFNL STKLELEKQV FGSRSLRPGP DSQRQDGRVH ASHQLTRVPA KPILPTATVR 1620  
 LPEMSTQSAS RYFVTSQSPR HWTNKPEITT YPSGALPENK QTTTFLSST TIPLPLHMSK 1680  
 PSIPSKPTDR RTDQNGVSK VFGNNNIPKA RNPVGKPPSP RHPHYSNGRL PFTTKTLSP 1740  
 POLGVTRRPQ IPTSPAPVR ERKVPGSYN RIHSHSTFHL DFGPPAPPLL HTPQKTGSPS 1800  
 TNLQNLPMVS STQSSISIT SVQSSGSEH QSSSKFPAGG PPASKFWSLG EKPQLITKSP 1860  
 QTVSVTARTD TVFPCBATGK PKPFVTWTKV STGALMTPTZ RIQRFEVLKV GTLVIRKQV 1920  
 QDRGQYMCCTA SNLHGLDRMV VLLSVTVQPP QILASHYQDV TVYLGDTIAM ECLAGTAPAP 1980  
 QISWIFPDRR VMQTVSPVES RITLHENRTL SIKEASPSDR GUYKCVASNA AGADSLAIRL 2040  
 HVAALPPIVH IETLENISLP PGLSIHICT AKAAPLPSVR WVLGDGTQIR PSQHLHGNLF 2100  
 VFENGTLVIR NLAPKDSGRY ECVAANLVGS ARRTVOLNVQ RAAANARITG TSPRTDVRV 2160  
 GGTLKLDCA SGPDPFRILM RLPSKRMIDA LPSFDSRIKV FANGTLVVKV VTDKADAGDYL 2220  
 CVARNKVGDD YVVLKVDVVM KPAKIEHKEE NDEKVFYGGD LKVDCAVATG LNPPEISWLP 2280  
 DGSVNSPMQ SDDSGGRTKR YVVFNNGLY FNEVGMRREG DYTCEAENQV GKDEMRVRV 2340  
 VVTAPATIRN KYTLAVQVYP GDVVTVACBA KGEPMKVTW LSPTNKVIPT SSEEKYIQD 2400  
 GTLLIQKQAR SSGSNYTCV RNSAGEDRKT VMLHVNQVPP KINGMNPIT TVREIAAGGS 2460  
 RKLIDCKAEG IPTPRVLNAP FEGVLPAPY YGNRITVHGN GSLDIRSLRK SDSVQLVOMA 2520  
 RNEGGEARLI VQLTVLEPME KPIFHDPISE KITAMAGHTI SLNCSAAGTP TPLSVWVLPN 2580  
 GTDLQSGQL QRFYHKAQDM LHSGLSSVD AGAYRCVARN AAGHTERLVS LKVGKPEAN 2640  
 KQYHNLVSI NBTLLKLLCP PPGAGQGRFS WTLPMGMHLE GPQTLGRVSL LDNGTLTVRE 2700  
 ASVFRDGYV CRMETEYGPS VTSIPVIVIA YPPRITSEPT FVIYTRPGNT VKLNCMAMGI 2760  
 PKADITWELP DKSHLKAQGV ARLYGNRFLH PQGSLTIQHA TQRDAGFYKC MAKNILGSDS 2820  
 KTTYIHVF

Seq ID NO: 418 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..5001

50 1 11 21 31 41 51  
 ATGCCAGGCA CAAACTAAC CGAACAGGC GCCCAGCAG ACTACAGAGT GATATTGAAG 60  
 ACCTCTCAAG AGGACGAATT GGATGTACCT GACGACATCA GCGTCCGGGT TATGTCTCAT 120  
 CAGTCTGTGC TTGTGTCTCT GGTGATCCT GTTCTGGAAA AACAGAAGAA AGTTGTGTGCA 180  
 55 TCAAGACAGT ACACCGTGCG CTATCGAGAG AAGGGGGAAT TGCCAGGTGT GGATTATAAG 240  
 CAGATCGCTA ACAGCGCTGT GCTGATTGAG AACCTGATTC CAGACACTGT GTATGAATTT 300  
 GCAGTCCGTA TTTCACAGGG TGAAAGAGAT GGCAATAGGA GTACGTCAGT CTTCCAAAGA 360  
 ACACAGAAAT CTCGCCCTCT CACAGCTCCT GAAACTCTGA ACGTCTGGCC AGTCAATGGC 420  
 AAACCTACAG TTGTGCTGTC ATCTTGGGAT GCGTACACAG AGACTGAGGG GAAAGTGAAA 480  
 60 GTCTGTCTGC TGGACACAGG ACTGTTTTC GTTCTCTCT TCCAACCAT TGCCAAATCA 540  
 TTTCAAGAATA CATTCTTTCA TACGCCCCGG CTCTCAAACC ATTTGGAGCA AAGTCCCTCA 600  
 CCTATCTCTG AGACACTACT TCTGCCCCGG TGGATGGTCT GCAGCCTGGG GAACGCTATC 660  
 TTTTCAAAAT CCGGGCCACA AACAGGAGAG GCCTGGGACC TCACTCCAAA GCCTTCATTG 720  
 65 TCCTATATCC AACAGAAATG CAGCTGTACC CAGAAGGATT TCAGTTGTCT AGCTTACCTG 780  
 ATGATATATC AAACCAAAAC AGTTAATAAA GATCCACAAC TGGAAAGGAG TGTTTTGG 840  
 CCATGTTTTC TTTTCTACTT CCTCACATT ATGCTGGATA TTGGCGGCTT TTCCTTCATT 900  
 ATGTGCTATG AAGACCCANN TGTTCCTCT TTGACAGGCA ATCTTTTAAA ATCTGTGTGCA 960  
 GCGATGAAG CAGATGTTCA GCAGAACACG GAGGACAAAT GGAAACCCGA AAAACCTGAG 1020  
 70 CCTTCTCAC CTCTCTCCAG AGCTCCAGCT TCCTCCCAAC ACCCTCTGT GCCTGCTTCT 1080  
 CCCCAGGGA GAAATGCCAA GGACCTCTT CTGACTTGA AGAACAAAAT ATTGGCTAAT 1140  
 GGTGGGGGCG CCGCAAAACC CAGCTTCCG GCCAAGAGG CAGAGGAGCT GGATCTTCAG 1200  
 TCGACAGAAA TCACTGGGGA GGAGGAGCTG GGTTCGCCGG AGGACTCGCC CATGTACCCC 1260  
 TCAGACACCC AAGACAGAAA ACGGACCTG AGGCGGCCAA GTAGACACGG CCACCTCGGTG 1320  
 75 GTTGCTCCCG GCAGGACTGC AGTGAGGGCC CGGATGCCAG CGCTGCCCGG AAGGGAAGGC 1380  
 GTAGATAAGC CTGGCTTTTC CTGGCCACG CAGCCCCGCC CAGGGGCGCC CCCCCTGGCT 1440  
 TCGGCTCTC CTGCCACCA CGGTCCACC CAGGGCACCT CTATCGTCC TTCCCTGCCT 1500  
 GCCAGCTTGA ATGACAAACA CTTGGTGGAC TCAGACGAAG ATGAGCGCGC TGTGGGCTCC 1560  
 TCCACCCCA AGGGGCGCTT CGCCAGCCCT TGTCCCCAG CGCCAGTCC 1620  
 80 CGCTCCAGCG TTCTCCGCGA CAGAAGCTCT GTGCACCCCG GCGCAAGGCC AGCTTCGCGG 1680  
 GCGCGGAGGA GAAATGCAAT AGGGGCGGCA GAGGAAGATT CCAATGCGCT AGCCCCACCC 1740  
 TCAAGACTTT TTCCACCCCA TGGGGGATCA TCTCGGCTGC TGCCACCCCA GCCACACCTG 1800  
 AGCTCTCCAC TTCTCAAGGG CCGGAAGGAT GGTGAGGACG CCCCAGCCAC CACTTCCAAT 1860  
 GCGCCATCAC GTTCCACCAT GTCTCTCTCC GTCTCTCTCC ATCTCTCTCC CAGGACGCG 1920  
 85 GTCTCTGAGG GAGCGGAGGC TTCTGATGGT GAAAGCCACG GTGACGCGCA TAGGGAAGAC 1980  
 GCGGGAAGGC AGCGGAGGC CACGCGCCAG ACGCTCGGG CCGGCGCTGC CTCTGGACAC 2040  
 TTCAATTGCT TCAGACCAA ACCCTTGTCT GCCAACGGGA GGTCTCCAAG CAGGTTGACG 2100  
 ATTGGGCGGG GACCTCGGCT GCAGCCCTCC AGCTCCCCAC AGTGGACTGT GCCCTCCGGA 2160

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65

GCGCCACCCCA GGGTTCCCTC TCACCTCTGAT TCCACACCTA AGCTTAGCTC AGGTATCCAT 2220  
 GGAGACGAGG AGGATGAGAA GCGCTTCTCT GCCACCGTTG TCAATGACCA CGTGCCTTCC 2280  
 TCCTCCAGGC AGCCCATCTC CCGGGGCTGG GAGGACTTAA GGAGAAGCCC GCAGAGAGGG 2340  
 GCCAGCCTGC ATCGGAAGGA ACCCATCCCA GAGAACCCTA AATCCACAGG GGCAGATACA 2400  
 CATCTCTAGG GCAAGTACTC CTCCTCGGCC TCCAAGGCTC AGGATGTTCA ACAGAGCACA 2460  
 GACGCGGACA CGGAGGGTCA TTCTCCCAAA GCACAGCCAG GGTCCACAGA CCGCCACGGG 2520  
 TCCCTTGTCT GTCTTCCCGC AGCACGGTCA CAGCAGCATC CCAGTGTTC CAGAAGGATG 2580  
 ACACCCCGCC GGGCCCCAGA ACAGCAGCCC CCTCTCCCGC TCGCCAGCTC CCAGCACCAC 2640  
 CCGGACCTCC AGAGCAGAGA CGCGGCTCGG TCACCTTCCC AGCCCAAGCT CTCACTGACC 2700  
 CAGGCGGGGC GGGCCCGCCC CAOGTCCGAG GCGCGCTCCC ACTCTCTCTC GGACCCTTAC 2760  
 ACGCGAGGCT CCAGAGGGAT GCTCCCCACG GCCCTCCAGA ACCAGGACGA GGATGCCACG 2820  
 GGCAGCTACG AOCAGCAGAG CACAGAAGTC GAGGCCACGG ATGTGCGGGC CCGCGCGCAC 2880  
 GCGCGCGGCG CCAAGGAGGC AGCTGCGTCC CTTCCTAAGC ACCAGCAGGT GGAGTCTCCC 2940  
 ACAGGCGCAG AGAGCAGTGG CGAACACAGG TCCAGCGCGC GACATGCGGC CTCCCCGCCC 3000  
 AGGCCCGAGC GACCCGCGGG CCCCAGTCCC CGCGCCCGGG TCCCCAGCAG GGCAGCGCGG 3060  
 GGGAGTGGCC AGCTCTCTTC CAAGCGGCCC CTGTCTTCCA AGTCCCAGCA GTCGGTCTCA 3120  
 GCGAGGAGCG AGGAGGAGGA GACGCGGGGG TTTTAAAG GCGGGAAGA AGACCTCTG 3180  
 TCTTCTCTG TGCCAAAGTG GCCCTCTTCC TCCACTCCCA GGGCGCGCAA AGACGCGGAT 3240  
 GGGAGCCTCG AGAGCAGAGA GAGGAGGCTT GCCATCGGCG GTGCTCTCGC CGGAGGGAGC 3300  
 CTGGCTCTG TGAGCGGACC TCTCCCCCA CCTCCAGGCA GCTCCCCCAG GGCTCTCCAC 3360  
 GTCCCTTCCC GACCGCGGCC TCGCAGGCTT GCCACCGTGA GCGCCGTCGC GGGCAGCCAC 3420  
 CCCTGGCGCG GGTACACCA CCGCGCCCGV CCTGGCCACT TCTCCACCAC CCGGATGCTG 3480  
 TCTTGGCGCC AGAGGATGAT GCATGCCAGA TTCCGTAAAC CTCTCTCCCG ACAGCCTGCC 3540  
 AGACCTCTTT ACAGACAAGG TTATAATGGC AGACCAATG TAGAAGGGAA AGTCTCTCT 3600  
 GGTAGTAATG GAAACCGAA TGACACAGAG ATTATCAATG GCCCTCAAGG AACAAAGTGG 3660  
 GTTGTGGACC GTTGTCTTGG GTTAGTATTG AATGCAGAG GAAGGTACCT CCAAGATTCA 3720  
 CATGGAATC CTCTTGGGAT TAACTAGGA GGAGATGGTC GAACCATGT AGATCTGGAA 3780  
 GGGACCCCGG TGGTGAATCC TGACGCGCTC CCACTCTTTG GGCAGGGGCG ACATGGCACA 3840  
 CCTCTGGCCA ATGCCAAGA TAAGCCAATT TTGAGTCTTG GAGGAAAGCC GCTGGTGGGC 3900  
 TTGGAGGTCA TCAAAAACAC CACCCATCCC CCTACCACTA CCATGCGGCC CACCACTACT 3960  
 ACAGCGCCCC TGCTTACCAC TACAACCCCG AGGCCACCA CTGCCACCA CATGCAGGCC 4020  
 ACCACTACTA CGACGCGCTT GCCTACCACT ACACCGAGGC CCACCACTGC CACCACCGCG 4080  
 OGCAGACCA CCAGCGCTCC AACCAACACA GTCCGAACCA CTACGCGGAC AACCAACCAC 4140  
 ACCACCCCCA AACCCACCAC TCCCATCCCC ACCTGTCCCC CTGGGACCTT GGAACCGCAC 4200  
 GACGATGATG GCAACCTGAT AATGAGTCCC AATGGGATCC CAGAGTGCTA CGCTGAAGAA 4260  
 GATGAGTCTT CGGCTTGGG GACTGACACT GCAGTACCTA CGGAAGAGGC CTACGTTATA 4320  
 TATGATGAAG ATTATGAATT TGAGACGTCA AGGCCACCAA CCACCACTGA GCCTTCGACC 4380  
 ACTGTACCA CACCGGCGGT GATCCGAGG GAGGCGCCA TCAGTTCCTT TCCTGAAGAA 4440  
 GAATTGATC TGGCTGAAG GAAACGATTT GTTGCTCCTT ACGTCACTA CCTAAATAAA 4500  
 GACCCATCAG CCGCTGTCTC TCTGACTGAT GCATGGATC ACTTCCAGT GGACAGCCTG 4560  
 GATGAATACA TCCCAATGA CCGTGAAGAG AGTGATCTGC CTCCCCAGCA TGCTCCCGC 4620  
 AACATCACCG TGGTGGCGGT GGAAGGTGTC CACTCATTTG TCAITGTGGA TTGGGACAAA 4680  
 GCCACCCAGG GAGATTGGGT CACAGGTTAT TTGGTTACA GTGCATCCTA TGAAGATTTC 4740  
 ATCAGGAACA AGTTTTCCAC TCAAGCTTCA TCAGTAACTC ACTTGCCCAT TGAGAACCTA 4800  
 AAGCCCAACA CGAGGTATTA TTTTAAAGTG CAAGCAGAAA ATCTCATAGG CTACGGACCT 4860  
 ATCAGCCCTT CGGCTCTCAT TGTCAACGAA TCAGATAATC CTCTGCTTGT TGTGAGGCC 4920  
 CAGGCGGGTG AGCTATCTGG ATCCCATTCG CTTTCAACAA TGATCCGAGC TACACGGACT 4980  
 GCCATGGAGG GCAATATGTG AAGCGCAGT GGTATCGAAA GTTCGTGGGA GTTGTCTT 5040  
 GTAATTCACT GAGGTATAAA ATCTACCTCA GTGACAACTT GAAAGATACA TTCTACGACA 5100  
 TTGGAGACAG CTGGGGAAGA GGTGAAGACC ATTGCCAATT TGTGATTCA CACCTTGATG 5160  
 GAAGAACAGG GCTCAGTCC TATGTAGAAG CCCTCCCTAC TATTCAAGGC TACTATGCGC 5220  
 AGTATCGTCA GGAGCCTGTC AGGTTTGGGA ACATCGGCTT CGGAACCCCC TACTACTATG 5280  
 TGGGCTGGTA CGAGTGTGG GTCTCCATCC CTGGAAGTGT GTAATCAGG GACCGTCATG 5340  
 CTGCAAGCTT GCGCTGCCCA GCCCCACCAA CTAAGTCCGA CTAGGGGCTG TGAGCAAGAA 5400  
 CAGCCAGCAT GCTCAGCCCC GCTGCCCTAG GTGCCAGGAA GGTCAACAGT GGACACTGGC 5460  
 CATCTCTGTC ATCTCAGTCT GGAACCTCAGT CCCACTTCTT GGCCTGGACA ATGAACAGGA 5520  
 TTCAGTTTGT CTGTAACTT TGCTTCTCTA CTTTCTTTTG TTTGTTTGTG ATAGCACATC 5580  
 CCAGAGACAT CAGAAACCGC CAACGTGATC AGTGTGATTT CCGAGACTTT TTAGGCATGA 5640  
 AATTGGACA CTTCACTATT TCCAGGAATA GCATATGCAC GCTGTTCTTG CTTTCATGAA 5700  
 TGCTACATGC TTTCTGTTTT TCTCATTTTG GATTCTTCCA AAACCTAAGT AATTAAAGCT 5760  
 TCAGTCTCCT TTGTATGCAG TAGAAAGGAA TTATTAATAA CACCACCAA GAAATAAAT 5820  
 ATATCTACT TGAAATTTAC TCTATGGACT TACCACATGC TAGAATAAAT GTATCAAATC 5880  
 TTATTGTAA ATTCTCAATT TTGATATATA TATGTATATA TGCATATACA TATCCACACT 5940  
 TGTCTGCAAG AATATTGATT AAAATTGCTA AATTGTACT TGTTCACCAA AAAAAA 6000

Seq ID NO: 419 Protein sequence  
Protein Accession #: Eos sequence

70  
75  
80  
85

1 11 21 31 41 51  
 MPGTLKTRTG APADYRVILK TSQEDELDPV DDISVRVMSS QSVLVSVDVP VLEKQKVVVA 60  
 SRQYTVRVRE KGLARWDYK QIANRRVLIE NLIPDTVYEF AVRISSQGERD GKWSTSVFQR 120  
 TPESAPTTAP ENLNVWVNG KPTVVAASWD ALPETEGKVK VCLLDLGLFS VSSFPQSAKS 180  
 FQNTFFHTPR LSNHLEQSPS PILETLLELPW WMVCSLGNAI FSKSGPQTGE AWDLTTPKPSL 240  
 SLCQEQSCST QKDFSLCLAYL IDIQTKQVNM DPQLEGSVFG PCFLFYFLTF MLDIGGFSFI 300  
 NCYEDPVSSL TGNLSKSVAA SKADVQQNTE DNGKPEKPEP SSPSPRAPAS SQHPSVPASP 360  
 QGRNAKDLLE DLNKLILANG GAPRKPQLRA KKAELDLQS TEITGZEEELG SREDSPMSPS 420  
 TDQDQKRLR PPSRHHGSVV APGRTAVRAR MPALPRRGV DKPGFSLATQ PRPGAPPSAS 480  
 ASPAHASTQ TSHRPSLPA SLNDNDLVDS DEDERAVGSL HPGKAPQPR PALSPSRQSP 540  
 SSVLRDRSSV HPGAKPASPA RRTPHSGAAE EDSSASAPPS RLSPPHGGSS RLLPTQPHLS 600  
 SPLSKGKGKDG EDAPATNSNA PSRSTMSSSV SSSLSSRTQV SEGAEASDGE SHGDGDREDG 660  
 GRQAEATAQT LRARPASGHF HLLRHKPFPA NGRSPSRPSI GRGPRLQPS SPQSTVPSRA 720  
 HPRVPSHSDS HPKLSSGIEG DEDEKPLFA TVVNDHVPSS SRQPISRGWE DLRRSPQRGA 780  
 SLHRKEPIEQ NPKSTGADTH PQGKYSSLAS KAQDVQQSTD ADTEGHSPKA QPGSTDRHAS 840  
 PARPPAARSQ QHPSVPRMT PGRAPEQQFP PFVATSQBHP GPQSRDAGRS PSQPRLSLTQ 900  
 AGRPRPTSQG RSHSSDPYFT ASSRGMPLTA LQNQDEDAQG SYDDDSSTEVE AQDVRAPAHA 960

ARAKEAAASL PKHQVESPT GAGAGGDHRS QRGHAAASPAR PSRPGGPQSR ARVPSRAAPG 1020  
 KSEPPSKRPL SSKSQSVSA EDEEEDAGF FKGGKEDLLS SSVKWPSSS TPRGGKADAG 1080  
 SLAKEEREPA IALAPRGGSL APVKRPLPPP FGSSPRASHV PSRPPPSAA TVSPVAGTHP 1140  
 WPRYTTTRAP GHFTTTPMLS LRQRMHARF RNPLSRQPAR PSYRQGYNGR PNVEGKVLPG 1200  
 SNGKPNQRI INGPQGTQKV VDLDRGLVLN AEGRYLQDSH GNPLRLIKLG DGRTIVDLEG 1260  
 TPVVSFDGLP LFGQGRHGT LANAQDKPIL SLGGKPLVLG EVIKKTTHPP TTTMQPTTTT 1320  
 TPLPTTTTFR PTTATTMQPT TTTTLPPTT PRPTTATRR TTTTPTTIV RTTTRTTTT 1380  
 TPKEPTTPIPT CPPGTLEHED DDGNLIMSSN GIPECTAEED EFSGLETDTA VPTTEAYVIY 1440  
 DEDYEFETSR PPTTTEPSTT ATTPRVIPEE GAISFPPEE FDLAGRKRFP APVVTYLNKD 1500  
 PSAPCSLTD LDFHFOVDSL EIIPNDLKKS DLPPQHAPRN ITVVAVEGCH SFVIVDWDKA 1560  
 TPGDLVTGYL VYSASYEDFI RNKFSQASS VTHLPIENLK PNTRYFKVQ AQNPHGYGPI 1620  
 SPVSFVTEG DNPLLVRFP GGRLSGSHSL SNMIPATRTA MDGNM

Seq ID NO: 420 DNA sequence  
 Nucleic Acid Accession #: NM\_022743  
 Coding sequence: 128..1237

1 11 21 31 41 51  
 GTGGATTTTA GAGATACCTC CCCTCCTTCT GCTCAGCTGC CTTCAGTAA TTAACCTCTT 60  
 TCTCTGCTGC AACACCCCTA CTGTCTCCG TGTATTGGCT TTTCTGGCA GCAGGAAGGA 120  
 AAAGCTGATG CGATGCTCTC AGTGGCCGCT GCGCAAATAC TGTAGTGCTA AGTGTGAGAA 180  
 AAAGCTTGG CCGAGCCACA AGCGGGAATG CAATGCTCTT AAAAGCTGCA AACCCAGATA 240  
 TCCTCCAGAC TCCGTTCGAC TTCTTGGCAG AGTTGTCTTC AAACCTATGG ATGGAGCACC 300  
 TTCAGATACA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAAAATATA ACAAACTGAC 360  
 TGAAGATAAG AAGAGGGGCC TCAGGCAACT CGTAATGACA TTTCACATT TCATGAGAGA 420  
 AGAATACAG GATGCTCTC AGCTGCCACC TGCCTTTGAC CTTTTGAAG CCTTTGCAAA 480  
 AGTGATCTGC AACTCTTTCA CCATCTGTAA TCGCGAGATG CAGGAAGTTG GTGTTGGCCT 540  
 ATATCCCATG ATCTCTTTGC TCAATCACAG CTGTGACCCC AACTGTTCGA TTGTGTTCAA 600  
 TGGGCCCCAC CTCTACTGCG GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT 660  
 CTGCTACTCG GATATGCTGA TGACCACTGA GGAGCCCGG AAGCAGCTGA GGGACCACTA 720  
 CTGCTTTGAA TGTACTGTCT TCCGTTGCCA AACCCAGGAC AAGGATGCTG ATATGCTAAC 780  
 TGGTGTAGAG CAAGTATGGA AGGAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA 840  
 GGCACACTGG AAGTGGGAGC AGGTCTGGC CATGTGCCAG GCGATCATAA GCAGCAATTC 900  
 TGAACGGCTT CCCGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGATGC 960  
 CTGCACTAAC CTCGGCTGT TGGAGGAAGC CTGTCTCTAT GGTACTCGGA CCATGGAGCC 1020  
 ATACAGGATT TTTTCCAG GAAGCCATCC CGTCAGAGGG GTTCAAGTGA TGAAGTTGG 1080  
 CAAACTGCAG CTACATCAAG GCATGTTTCC CCAAGCAATG AAGAATCTGA GACTGGCTTT 1140  
 TGATATTATG AGAGTGACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT 1200  
 TTTAAGAAAG TGGCAGCCCA ACATCAGAGC ATCTAAGGG AACGCAGTCA GAGGGAATA 1260  
 CGCGTGTGT CTGTGTGAA TGCCCTATTG AGGTACACA CTCTATGCTT TGTATGCTGT 1320  
 GTGAACCTCT CTATTGGAAT ATTCGTGTC GTGTGTGTGT AGGTAAATA AGGCAGACAT 1380  
 GGTGTGCAA CCACAGAAT CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAAC 1440  
 ATTTGGTTGA GGATGCCAAA AAAAAAAAAA AAAAAA

Seq ID NO: 421 Protein sequence  
 Protein Accession #: NP\_073580

1 11 21 31 41 51  
 MRCSQCRVAK YCSAKQKKA WPDHKRECK LKSCKPRYP DSVRLLRGV FKLMDGAPSE 60  
 SEKLYSFYDL ESNINKLTED KKEGLRLQVM TPQHFMRBEI QDASQLPPAF DLFEAPAKVI 120  
 CNSFTICNAE MQEVGVGLYP SISLLNHSCD PNCISIVFNGP HLLLRVRDI EVGEELTICY 180  
 LDMIMTSEER RKQIRDQYCF ECDCFRCQTO DKADMLTGD EQVMKEVORS LKKIEELKAH 240  
 WKWEQVLAMC QAISSNSER LPDINIYQLK VLDCAADACI NLGLLEBALF YGTRTMEPYR 300  
 IPPPGSHFVR GVQVMKVGKL QLEQGMFPQA MNLRLLAFDI MRVTHGREHS LIEDLLILLE 360  
 ECDANIRAS

Seq ID NO: 422 DNA sequence  
 Nucleic Acid Accession #: NM\_003014.2  
 Coding sequence: 238..648

1 11 21 31 41 51  
 GCGCGGTTCG CGCCCGAAG GCTGAGAGCT GCGCTGCTC GTGCCCTGTG TGCCAGACGG 60  
 CGGAGCTCCG CGGCCGACC CCGCGGCCCC GCTTTGCTGC CGACTGGAGT TTGGGGGAAG 120  
 AAATCTCTCT GCGCCCGAGA AGATTTCTTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT 180  
 GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTCCGAG CGCGAGAGGG CAGTGCCATG 240  
 TTCTCTCCA TCCTAGTGGC GCTGTGCTGC TGGCTGCACC TGGCGCTGGG CGTGGCGGCG 300  
 GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGCGACA TGCCCTGGAA CATCAAGCGG 360  
 ATGCCCAACC ACCTGCACCA CAGCACGCG GAGAACGCCA TCCTGGCCAT CGAGCAGTAC 420  
 GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTCTGT TGCCATGTAC 480  
 GCGCCCATTT GCACCTTGA GTTCTGTCAC GACCCATATCA AGCCGTGCAA GTCGGTGTGC 540  
 CAAAGCGCGC GCGACGACTG CGAGCCCTCT ATGAAGATGT ACAACCAAG CTGGCCCGAA 600  
 AGCCTGGCCT GCGACGAGCT GCCTGTCTAT GACCGTGGCG TGTGCATTTT GCCTGAAGCC 660  
 ATCGTCAAGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACAGCA CATGATGGTA 720  
 CAGGAAGGCG CTCTTGATGT TGACTGTAAA CGCCTAAGCC CCGATCGGTG CAAGTGTAAA 780  
 AAGGTGAAGC CAACCTTGGC AACGTATCTC AGCAAAACT ACAGCTATGT TATTATGCC 840  
 AAAATAAAG CTGTGCAAG GAGTGGCTGC AATGAGGTCA CAACGGTGGT GGATGTAAAA 900  
 GAGATCTTCA AGTCTCTATC ACCCATCCCT CGAATCAAG TCCCGCTCAT TACAAATTCT 960  
 TCTTGCCAGT GTCCACACAT CCTGCCCAT CAAGATGTTT TCATCATGTG TTACGAGTGG 1020  
 CGTTCAAGGA TGATGCTTCT TGAATAATGC TTAGTTGAAA AATGGAGAGA TCAGCTTAGT 1080  
 AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAAG 1140  
 AAGAAAACAG CCGGGCGCAC CAGTCGTAGT AATCCCCCA AACCAAGGG AAAGCCTCCT 1200  
 GCTCCCAAC CAGCAGTCC CAAGAAGAAC ATTAATACTA GGAGTGCCCA GAAGAGAAAC 1260  
 AACCAGAAA GAGTGTAGC TAACTAGTTT CCAAGCGGA GACTTCGAC TTCTTACAG 1320  
 GATGAGCTG GGCATTGCTT GGCAGACCTT ATGTAAGGCC ATGTGCCCT TGCCCTAACA 1380

ACTCACTGCA GTGCTCTCA TAGACACATC TTGCAGCATT TTCTTAAGG CTATGCTTCA 1440  
 GTTTTCTTT GTAAGCCATC ACAAGCCATA GTGCTAGGTT TGCCCTTTGG TACAGAAGGT 1500  
 GAGTTAAAGC TGGTGGAAAA GGCTTATTGC ATTGCATTCA GAGTAACCTG TGTGCATACT 1560  
 CTAGAAGAGT AGGGAAAAATA ATGCTTGTGA CAATTGCGACC TAATATGTGC ATTGTAAAAAT 1620  
 AAATGCCATA TTTCACAAAC AAACAGTAAT TTTTTCACAG TATGTTTTAT TACCTTTTGA 1680  
 TATCTGTGTG TGCATGTGTA GTGATGTITT AAAATGTGAT GAAATATATA TGTTTTAAAG 1740  
 AAGGAACAGT AGTGGAAATGA ATGTTAAAG ATCTTTATGT GTTTATGGTC TGCAGAAGGA 1800  
 TTTTGTGAT GAAAGGGGAT TTTTGAAGAA ATTAGAGAAG TAGCATATGG AAAATTATAA 1860  
 TGTGTTTTT TACCAATGAC TTCAGTTTCT GTTTTAGCT AGAAACTTAA AAACAAAAAT 1920  
 AATAATAAAG AAAATAAAT AAAAAGGAGA GGCAGACAAT GTCTGGATTC CTGTTTTTGT 1980  
 GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTTAAGC AGCACCAGAA 2040  
 ACAGTGAGTT TGTCTGTACC ATTAGGAGTT AGTACTAAT TAGTTGGCTA ATGCTCAAGT 2100  
 ATTTTATACC CACAAGAGAG GTATGTCAC CTCTTACTT CCCAGGACAT CCACCCTGAG 2160  
 AATAATTGGA CAAGCTTAAA AATGGCCTTC ATGTGAGTGC CAAATTTTGT TTTTCTTCAT 2220  
 TTAATATATT TCTTTGCCATA AATACATGTG AGAGGAGTTA AATATAAATG TACAGAGAGG 2280  
 AAAGTTGAGT TCCACCTCTG AATGAGAAT TACTTGACAG TTGGGATACT TTAATCAGAA 2340  
 AAAAGAACT TATTGTCAGC ATTTTATCAA CAAATTTTCT AATTGTTGAC AATTGGAGGC 2400  
 ATTTATTTTA AAAACAATT TTATTGSCCT TTTGCTAACA CAGTAAGCAT GTATTTTATA 2460  
 AGGCATTCAA TAAATGCACA ACGCCCAAAG GAAATAAAT CCTATCTAAT CCTACTCTCC 2520  
 ACTACACAGA GGTAACTCAT ATTAGTATT TGGCATATTA TTCTCCAGGT GTTTGCTTAT 2580  
 GCACTTATAA AATGATGTA ACAAAATAAA TAGGAACCT GTATACATGT GTTTCATAAC 2640  
 CTGCTCTCTT TGCTGGGCC TTATTGAGA TAAGTTTCC TGTCAAGAAA GCAGAAACCA 2700  
 TCTCATTTCT AACAGCTGTG TTATATTCCA TAGTATGCAT TACTCAACAA ACTGTTGTGC 2760  
 TATTGTATAC TTAGTGTGTT TCTTCACTGA CAATCTGAA TAAACATCTC ACCGGAATTC

Seq ID NO: 423 Protein sequence  
 Protein Accession #: NP\_003005.1

1 11 21 31 41 51  
 MFLSILVALC LNLHLALGVR GAPCEAVRIP MCRHMPWNIT RMPNHLHST QENAILAIEQ 60  
 YEELVDVNVK AVLRFFFCAM YAPICTLEFL HDPIKPKSV QQRARDCEP LMKMYNHSWP 120  
 ESLACDELFP YDRGVCISPE AIVTDLPELV KWIDITPDMV VQERFLDVDC KRLSPDRCKC 180  
 KKVKPTLATY LSKNYSYVIH AKIKAVQSRG CNEVTVVDV KEIPKSSSPI PRTOVPLITN 240  
 SSCQCFHLLP HQDLVIMCYE WSRMMLLEN CLVEKWRDQL SKRSIQWEER LQEQRTVQD 300  
 KKTIAGRTSR SNPKPKGKPK PAPKPSFKK NIKTRSAQKR TNPKRK

Seq ID NO: 424 DNA sequence  
 Nucleic Acid Accession #: BC010423  
 Coding sequence: 248..1780

1 11 21 31 41 51  
 CACAGCGTGG GAAGCAGCTC TGGGGGAGCT OGGAGCTCCC GATCACGSGT TCTTGGGGGT 60  
 AGCTACGCGT GGGTGTGTAG AACGGGGCCG GGGCTGGGGC TGGGTCCCTT AGTGGAGACC 120  
 CAAGTGCAGG AGGCAAGAAC TCTGCAGCTT CCTGCCCTTC GGGTCAGTTC CTTATTCAAG 180  
 TCTGCAGCGC GCTCCCGAGG AGATCTCGGT GGAACCTCAG AAACGCTGGG CAGTCTGCCT 240  
 TTCAACCATG CCGCTGTCTCC TGGGAGCCGA GATGTGGGGG CCTGAGGCTT GGGTCTGTCT 300  
 GCTGTCTACTG CTGGCATCAT TTACAGGCGG GTGCCCCGGG GGTGAGCTGG AGACCTCAGA 360  
 CGTGGTAACT GTGGTGTCTG GCCAGGACGC AAAACTGCCC TGCTTCTACC GAGGGGACTC 420  
 CGGCGAGCAA GTGGGGCAAG TGGCATGGGC TCGGTGGAGC GGGGGCGAAG GCGCCAGGA 480  
 ACTAGCGCTA CTGCACTCCA AATACGGGCT TCATGTGAGC CCGCTTACG AGGGCCGCGT 540  
 GGAGCAGCGT CCGCCCCCAT GCAACCCCTT GGAACGCTCA GTGCTCCTGC GCAACGCACT 600  
 GCAGGCGGAT GAGGGCGAGT ACGAGTGCAG GTTCAGCACC TTCCCCCGCG GCAGCTTCCA 660  
 GGGCGGGCTG CGGCTCCGAG TGCTGGTGCC TCCCTGCCC TCACGTGAATC CTGTTCCAGC 720  
 ACTAGAAGAG GGCCAGGGCC TGACCTTGGC AGCCTCTGCG ACAGCTGAGG GCAGCCCAGC 780  
 CCCACGCGTG ACTTGGGACA CGGAGGTCAA AGGCACAAAG TCCAGCGGTT CCTTCAAGCA 840  
 CTCCGCTCTC GCTGCGCTGA CCTCAGAGTT CCACTTGGTG CCTAGCCGCA GCATGAATGG 900  
 GCAGCCCATG ACTGTGTGTT TGTCCTATCC TGGCCTGCTC CAGGACCAAA GGATCACCCA 960  
 CATCTCCACG GTGTCTCTTC TTGCTGAGGC CTCTGTGAGG GGCCTTGAAG ACCAAAATCT 1020  
 GTGGCACATT GGCAGAGAAG GAGCTATGCT CAAATGCGCT AGTGAAGGGC AGCCCCCTCC 1080  
 CTCATACAACT TGACACCGGC TGGATGGGCC TCTGCCAGT GGGGTACGAG TGGATGGGGA 1140  
 CACTTTGGGC TTTCCTCCAC TGACCACTGA GCACAGCGGC ATCTACGTCT GCCATGTGAC 1200  
 CAATGAGTTC TTCTCAAGGG ATTCTCAGGT CACTGTGGAT GTTCTTGACC CCCAGGAAGA 1260  
 CTCTGGGAAG CAGGTGGACC TAGTGTGAGC CTGGTGGGTG GTGGTGGGTG TGATCGCCGC 1320  
 ACTCTGTGTC TGCCCTCTGG TGGTGGTGGT GGTGCTCATG TCCGATACCC ATCGGCGCAA 1380  
 GGCCACGAGC ATGACCCAGA AATATGAGGA GGAGCTGACC CTGACCGAGG AGAACTCCAT 1440  
 CCGGAGGCTG CATTTCCATC ACACGGACCC CAGGAGCCAG CCGGAGGAGA GTGTAGGGCT 1500  
 GAGAGCCGAG GGCACCCCTG ATAGTCTCAA GGACAAAGT AGCTGCTCTG TGATGAGTGA 1560  
 AGAGCCCGAG GGCAGCAGT ACTCCAAGCT GACCAAGGTG AGGGAGATAG AAACACAGAC 1620  
 TGAAGTCTG TCTCCAGGCT CTGGGCGGGC CGAGGAGGAG GAAGATCAGG ATGAAGGCAT 1680  
 CAAACAGGCC ATGAACCAAT TTGTTTCAGGA GAATGGGACC CTACGGGCCA AGCCCCAGGG 1740  
 CAATGGCATC TACATCAATG GGCAGGGACA CCTGCTCTGA CCCAGGCGCT CCTCCCTTCC 1800  
 CTAGGCGCTG CTCTTCTGTG TGACATGGGA GATTTTAGCT CATCTTGGGG GCCTCCTTAA 1860  
 ACACCCCATC TTCTTGGGGA AGATGCTCCC CATCCACTG ACTGCTTGAC CTTTAACTCC 1920  
 AACCCCTCTG TTCTCAGGGA GGGCTCCACC AATTGAGTCT CTCCACCAT GCATGCAGGT 1980  
 CACTGTGTGT GTGCATGTGT GCCTGTGTGA GTGTGACTG ACTGTGTGTG TGTGGAGGGG 2040  
 TGACTGTCCG TGGAGGGGTG ACTGTGTCCG TGGTGTGTAT TATGCTGTCA TATCAGAGTC 2100  
 AAGTGAACCTG TGGTGTATGT GCCACGGGAT TTGAGTGTGT GCGTGGGCAA CACTGTGAGG 2160  
 GTTGTGGGTG TGTGTGATGT GGCTGTGTGT GACCTCTGCC TGAAGAAAGCA GGTATTTTCT 2220  
 CAGACCCAGG AGCAGTATTA ATGATGCAGA GGTGGAGGA GAGAGGTGGA GACTGTGGCT 2280  
 CAGACCCAGG TGTGGGGGCA TAGCTGGAGC TGGAACTCTG CTCCGGTGTG AGGGAACCTG 2340  
 TCTCTTACCA CTTGGGAGCC ATGGGGGCAA GTGTGAAGCA GCGAGTCCCT GGGTCAGCCA 2400  
 GAGGCTTGAA CTGTACAGA AGCCCTCTGC CCTCTGGTGG CCTCTGGGCG TGCTGCATGT 2460  
 ACATATTTTC TGTAAATATA CATGCGCGGG GAGCTTCTTG CAGGAATACT GCTCCGAATC 2520  
 ACTTTTAATT TTTTCTTTT TTTTCTCTG CCCTTTCAT TAGTTGTATT TTTTATTAT 2580  
 TTTTATTTT ATTTTCTTTT AGAGTTTGAG TCCAGCCTGG ACGATATAGC CAGACCCCTG 2640



CTGTAAAAAA ACCAAAAACC AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 425 Protein sequence  
Protein Accession #: AAH10423

1 11 21 31 41 51  
MPLSLGAEMW GPEAWLLLLL LLASFTGRCP AGELETSDDV TVVLGQDAKL PCFYRGDSGE 60  
QVQGVAVARV DAGEGAQELA LLHSKYGLEV SPAYEGRVEQ PPPPNPLDG SVLLRNVAQA 120  
DEGEYECRVS TFPAGSFQAR LRLRVLPPL PSINPGPALE EGGLTLAAS CTABGSPAPS 180  
VTWDETVKGT TSSRSPKHSR SAAVTSEFHL VPSRSMNGQP LTCVVSHPGL LQDQRITHIL 240  
HVSFLAEASV RGLDQNLWH IREGAMLKC LSEGQPPPSY NWTRLDGPLP SGVRVDGDTL 300  
GFPPLTTEHS GIYVCHVNE FSSRDSQVTV DVLDPQEDSG KQVDLVASV VVGVIAALL 360  
FCLLVVVVL MSRYHRRKAQ QMTQKYEEL TLTRENSIRR LHSHTDPRS QPEESVGLRA 420  
EGHPDSDLKN SSCSVMSSEF EGRSYSTLT VREIETQTEL LSPGSGRAEB EEDQDEGIKQ 480  
AMNHPVQENG TLRAKPTNG IYINGRGHLV

Seq ID NO: 426 DNA sequence  
Nucleic Acid Accession #: NM\_003474.2  
Coding sequence: 37..3036

1 11 21 31 41 51  
CACTAAAGCT CTTCCTAGTC CCGGGGCCAA CTCGGACAGT TTGCTCATT ATTGCAACGG 60  
TCAAGGCTGG CTGTGCGCAC AACGCGCGCG CGCGGACGCA CGCACACACA CGGGGGGAAA 120  
CTTTTAAATA AATGAAAGGC TAGAAGAGCT CAGCGCGCGG CGGGGCGGTG CGCGAGGCT 180  
CGGAGCTGTA CTGCGCGAGG CAGGAAATCC CTCGCTGCG GACGCGCGGC CCCGCTCGGC 240  
GCCCGCGTGG GATGCTGCAG CGCTCGCGCG CGGCGCGGAG AGCTGCTGCA CTGAAGGCGG 300  
GCGAGCATGG CAGCGCGCGC GCTGCGCGTG TCCCGCGCCC GCGCGCTCT GCTGCGCTG 360  
GCTGATGAAG TTGTCACTGC CTCTGTTCGG AGTGGGACC TCTGATCCC AGTGAAGAGC 420  
TTGACTCCA AGAATCATCC AGAAGTGTG AATATTGAC TACAACGGGA AAGCAAGAA 540  
CTGATCATA ATCTCGAAG AAATGAAGT CTCAITGCCA GCAGTTTAC GGAAACCCAC 600  
TATCTGCAAG ACGGTACTGA TGCTCTCCTC GCTCGAAT ACACGTAAT TCTGGGTAC 660  
TGTTACTACC ATGGCATGT ACGGGGATAT TCTGATCAG CAGTCAGTCT CAGCACGTGT 720  
TCTGGTCTCA GGGGCTTAT TGTGTTGAA AATGAAAGCT ATGTCTTGA ACCAATGAAA 780  
AGTGCAACCA ACAGATACAA ACTCTCCCA GCGAAGAGC TGAAGAGCGT CCGGGGATCA 840  
TGTGATCAC ATCAACACAC ACCAAACCTC GCTGCAAGA ATGTGTTCC ACCACCTCT 900  
CAGACATGGG CAAGAAGGCA TAAAGAGAG ACCCTCAAG CAACTAAGTA TGTGGAGCTG 960  
GTGATCGTGG CAGACAACCG AGAGTTTCAG AGGCAAGGAA AAGATCTGGA AAAAGTTAAG 1020  
CAGCGATTAA TGAGATTGC TAATCACTG GACAAGTTT ACAGACCAT GAACATTGG 1080  
ATCGTGTGG TAGGCTGGA AGTGTGGAAT GACATGGACA AATGCTCTGT AAGTCAGGAC 1140  
CCATTACCA GCTCTCATGA ATTTCTGGAC TGGAGGAAGA TGAAGCTCT ACCTCGCAA 1200  
TCCCATGACA ATGCGCAGCT TGTCAGTGGG GTTATTTC C AAGGACACAT CATCGGCATG 1260  
GCCCAATCA TGAGCATGTG CACGGCAGAC CAGTCGCGG GAATTGTCTAT GGACCATCA 1320  
GACAATCCCC TTGGTGCAGC CGTGACCTCG GCACATGAGC TGGGCCACAA TTTCGGGATG 1380  
AATCATGACA CACTGGACAG GGGCTGTAGC TGTCAAATGG CGGTTGAGAA AGGAGGCTGC 1440  
ATCATAGCG CTTCACCGG GTACCCATT CCATGTTGT TCAGCAGTTG CAGCAGGAAG 1500  
GACTTGAGA CCAAGCTTGA GAAAGGAATG GGGGTGTGCC TGTTTAACT GCGGGAAGTC 1560  
AGGGAGTCTT TCGGGGCCCA GAAGTGTGG AACAGATTG TGAAGAAGG AGAGAGTGT 1620  
GACTGTGGGG AGCCAGAGGA ATGTATGAAT CGCTGCTGCA ATGCCACCAC CTGTACCCTG 1680  
AAGCGGACAG CTGTGTGCGC ACATGGGCTG TGCTGTGAAG ACTGCCAGCT GAAGCCTGCA 1740  
GGAAACAGGT GCAAGGATC CAGCAACTCC TGTGACCTCC CAGAGTTCTG CACAGGGGCC 1800  
AGCCCTCACT GCCAGGCCAA CGTGTACCTG CACGATGGGC ACTCATGTCA GGATGTGGAC 1860  
GGCTACTGCT ACAATGGCAT CTGCCAGACT CACGAGCAGC AGTGTGTAC ACTCTGGGGA 1920  
CCAGGTGCTA AACCTGCCCC TGGGATCTGC TTTGAGAGAG TCAATTCTGC AGGTGATCCT 1980  
TATGGCACT GTGGCAAGT CTCGAAGAGT TCCTTTGCCA AATGCGAGAT GAGAGATGCT 2040  
AAATGTGAAA AATCTCAGTG TCAAGGAGGT CCGACGCGGC CAGTCATTGG TACCAATGCC 2100  
GTTTCCATAG AAACAAACAT CCCCCTGCAG CAAGGAGGCC GGATTCTGTG CCGGGGAC 2160  
CACGTGTACT TGGCGCATGA CATGCGGAC CCAGGGCTTG TGCTGCAAG CACAAGTGT 2220  
GCAGATGGA AAATCTGCCT GAATCTGCAA TGTCAAATA TTAGTGTCTT TGGGTTTAC 2280  
GAGTGTGCAA TGCAAGTCCA CGGCAGAGGG GTGTGCAACA ACAGGAAGAA CTGCCACTGC 2340  
GAGGCCCACT GGCACCTCC CTCTGTGAC AAGTTTGGCT TTGGAGGAAG CACAGACAGC 2400  
GGCCCCACT GGCAGCAGA TAACCAAGGT TTAACCATAG GAATTCTGTT GACCATCTGT 2460  
TGTCTTCTTG CTGCGGATT TGTGTTTAT CTCAAAAGGA AGACCTTGAT ACGACTGCTG 2520  
TTTACAAATA AGAAGACCAC CATTGAAAAA CTAAGGTGTG TGGCGCCTTC CCGGCCACCC 2580  
CGTGGCTTCC AACCTGTCA GGCTCACCTC GGCACCTTG GAAAAGGCCT GATGAGGAAG 2640  
CCGCCAGATT CCTACCCACC GAAGGACAA CCCAGGAGAT TGCTGCAGTG TCAGAAATGT 2700  
GACATCAGCA GACCCCTCAA CGGCCTGAAT GTCCCTCAGC CCGAGTCAAC TCAGCGAGTG 2760  
CTTCTTCCCC TCCACCGGGC CCCACGTGCA CCTAGCTGCC CTGCCAGACC CCGCCAGGCC 2820  
AAGCCTGCAC TTAGGCAGGC CCAGGGGACC TGTAAAGCAA ACCCCCTCA GAAGCCTCTG 2880  
CCTGCAGATC CTCTGCGCAG AACCACTCGG CTCACCTAG CCTTGGCCAG GACCCAGGA 2940  
CAATGGGAGA CTGGGCTCCG CCTGGCACCC CTCAGACCTG CTCACAATA TCCACACCA 3000  
GTGCCAGAT CACCCACAC CGCCTATATT AAGTGAGAAG CCGACACCTT TTTTCAACAG 3060  
TGAAGACAGA AGTTTGCACT ATCTTTGAGC TCCAGTTGGA GTTTTGTGTA CCAACTTTTA 3120  
GGATTTTTTT TAATGTTTAA AACATCATT CTATAAGAAC TTTGAGTAC TGCCTCAGT 3180  
GCTGTGCTGT GCTATGGTGC TCTGTCTACT TGCACAGSTA CTGTAAATT ATTAATTAT 3240  
GCAGATGTT GATTACAGTG CAGTGCCTG TAGTAGGCAT TTTTACCATC ACTGAGTTT 3300  
CCATGGCAGG AAGGCTTGT GTGCTTTAG TATTTAGT AACTTGAAT ATCTGCTTG 3360  
ATGGGATTCT GGCAGGATG TGTGTTCTT CTGATCAAG CTTATTGGA AAGCAGTCCC 3420  
CCAACTACCC CCACTGTGTC TTAGGTGACC AGATGAGCT CAAGAGATCC CAGTAGAAT 3480  
CTCAGTTGAT TTTCTGGAAT CCCCCTCA GGCAGAGCC AAGGGCTTC AGGTCCAGGC 3540  
TGTGTTGGC TTTCAGGAG GCCCTGTGCC CCTTGACAAC TGGCAGGCG GCTCCAGGG 3600  
ACACCTGGGA GAAATCTGCG TTTGCGCAG GAAGCTTGG TGAAGACCTG GGTGTCAGAC 3660  
AGGAATCTTA AGGTGTAGCC ACACAGGAT AGAGACTGGA ACCTAGACA AGCCAGAACT 3720  
TGACCCCTGAG CTGACCAAGC GTGAGCATGT TTGGAAGGG TCTGTAGTGT CACTCAAGGC 3780  
GGTCTGTGAT AGAATGCCA AGCACTTCT TTTCTGCTG TCTTTCTAG AGCACTGCCA 3840



CCAGTAGGTT ATTTAGCTTG GGAAAGGTGG TGTTCCTGTA AGAAACCTAC TGCCCAAGCA 3900  
 CTGCAAAACCG CCACCTCCCT ATACTGCTTG GAGCTGAGCA AATCACCACA AACTGTATA 3960  
 CAATGATCCT GTATTTCAGAC AGATGAGGAC TTTCATGGG ACCACAACCTA TTTTCAGATG 4020  
 5 TGAACCACTA ACCAGATCTA GTCAATCAAG TCTGTTTACT GCAAGGTTCA ACTTATTAA 4080  
 AATTAGGCAG ACTCTTTATG CTTCGAAAAA CTACAACCAA TGAATGTGA TGTTTCATGGG 4140  
 TATAGTTCAT GTCTGCTATC ATTAITCGTA GATATTGGAC AAAGAACCTT CTCTATGGGG 4200  
 CATCCTCTTT TTCCAACCTG GCTGCAGGAA TCTTTAAAG ATGCTTTTAA CAGAGTCTGA 4260  
 ACCTATTCTT TAAACACTTG CAACCTACCT GTTGAGCATC ACAGAATGTG ATAAGGAAAT 4320  
 10 CAACTTGCTT ATCAACTTCC TAAATATTAT GAGATGTGGC TTGGGCAGCA TCCCCTTGAA 4380  
 CTCTTCACTC TTCAAATGCC TGAATAGGGA GCCATGTTC ACAGGTCTT TAAAGTGACT 4440  
 AATGGCATGA GAAATACAAA AATACTCAGA TAAGGTAAAA TGCCATGATG CCTCTGCTT 4500  
 CTGGACTGGT TTTACACATA GAAGACAATT GACAACAGTT ACATAATTC CTCTGAGTGT 4560  
 TTTATGAGAA AGCCTTCTTT TGGGGTCAAC AGTTTTCTTA TGCTTTGAAA CAGAAAAATA 4620  
 15 TGTACCAAGA ATCTTGGTTT GCCTTCCAGA AAACAAAAC GCATTTCAT TCCCGGTGT 4680  
 TCCCACGTGT ATCTAGGCAA CATAGTATTC ATGACTATGG ATAACTAAA CACGTGACAC 4740  
 AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTCGTATAGC ATGCATCTGT 4800  
 TTATTCTATA GTTATTAAAT TCTTTAAAT GTAAAGCCAT GCTGGAAAAT AATACTGCTG 4860  
 AGATACATAC AGAATTACTG TAACTGATTA CACTTGGTAA TTGTACTAAA GCCAAACATA 4920  
 20 TATATACTAT TAAAAAGGTT TACAGAATT TATGGTGAT TACGTGGGCA TTGCTTTTTT 4980  
 AGATGCCCAA ATCCTTAGAT CTGGCATGTT AGCCCTTCTT CCAATTATAA GAGGATATGA 5040  
 ACCAAAAAA AAAAAAAA AA

Seq ID NO: 427 Protein sequence

Protein Accession #: NP\_003465

25 1 11 21 31 41 51  
 MAARPLFPVP ARALLLALAG ALLAPCEARG VSLWNEGRAD EVVSASVRSQ DLWIPVKSFD 60  
 30 SKNHPEVLNI RLQRESKELI INLERNEGLI ASSFTETHYL QDGTDVSLAR NYTVILGHY 120  
 YHGHVGRGSD SAVSLSTCSG LRGLIVFENE SYVLEPMKSA TNRYKLPFAK KLSKVRGSCG 180  
 SHHNPINLAA KNVFPFPPSP WARRHKRETL KATKYVELVI VADNREFQRO KDKLEKVKQR 240  
 LIBIANHVDK FYRPLNIRIV LVGVEVWDM DKCSVSQDPF TSLHEFLDWR KMKLLPRKSH 300  
 DLAQLVSGVY FQGTITGMAP IMSMCTADQS GGIVMDHSDN PLGAAVTLAH ELGHNFGMNH 360  
 35 DTLDRGSCQ MAVEGGCCIN NASTGYPPFM VFSSCSRKDL ETSLEKMGV CLFNLEPVRE 420  
 SFGGQKCNR FVEEGEECD GEPEECMNR CNATTTCTLK DAVCAHGLCC EDCQLKPGT 480  
 ACRDSNSCD LPEFCTGASP HCPANVYLHD GHSQDQVDY CYNGICQTHE QQCCTLWGP 540  
 AKPAPGICFE RVNSAGDPYG NCGKVSKESS AKCEMRDAK GKIQCGGAS RPVIGTNAV 600  
 IETNIPLQGG GRILCRGTHV YLGDDMPDPG LVLAGTKCAD GKICLNRCQ NISVFGVHEC 660  
 40 AMQCHRGVC NNRKNCHCEA HWAPPFCDKF GFGGSTDGSP IRQADNQLT IGLVTLILCL 720  
 LAAGFVVYLK RKLTRLRLFT NKKTITIEKL CVRPSRPPRG FQPCQAHLG LKGLMRKPP 780  
 DSYFPKDNPR RLLQCNVDI SRPLNGLNVP QPQSTQRLV PLHRAPRAPS VPARPLPAK 840  
 ALRQAGGTCK PNFPQKPLPA DPLARTTRLT HALARTPGQW ETGLRLAPLR PAPQYHQVP 900  
 RSTHTAYIK

Seq ID NO: 428 DNA sequence

Nucleic Acid Accession #: NM\_003714

Coding sequence: 135..1043

50 1 11 21 31 41 51  
 GAGGAGGAGG GAAAAGGCGA GCAAAAAGGA AGAGTGGGAG GAGGAGGGGA AGCGGCGAAG 60  
 GAGGAAGAGG AGGAGGAGGA AGAGGGGAGC ACAAGGATC CAGGTCTCCC GACGGGAGGT 120  
 TAATACCAAG AACCATGTGT GCCGAGCGGC TGGGCCAGTT CATGACCCGT GCTTTGGTGT 180  
 55 TGCCACCTT TGACCCGCGC CGGGGACCG ACGCCACCAA CCCACCGAG GTTCCCAAG 240  
 ACAGGAGCTC CCAGCAGAAA GCGCGCTGT CCTTCAGAA TACAGCGAG ATCCAGCACT 300  
 GTTGGTCAA CGCTGGCAT GTGGGTGTG GCGTGTGTA ATGTTTCAG AACACTCTT 360  
 GTGAGATTG GGGCTTACAT GGGATTGCA TGACTTTCT GCACACGCT GGAATAATTG 420  
 ATGCCAGGG CAAGTCATT ATCAAGACG CCTTGAATG TAAGGCCAC GCTCTGCGC 480  
 ACAGGTTGG CTGCATAAG CGGAAGTGC CGGCCATCAG GGAATGGTG TCCAGTTGC 540  
 60 AGCGGGAATG CTACCTCAAG CAGACCTGT GCGCGCTGC CCAGGAGAAC ACCCGGTGA 600  
 TAGTGGAGAT GATCCATTTC AAGGACTTGC TGCTGCAGA ACCCTACGT GACCTCTGA 660  
 ACTTGTCTCT GACCTGTGGG GAGGAGGTGA AGGAGGCCAT CACCCACAGC GTGCAGGTTC 720  
 AGTGTGAGCA GAACCTGGGA AGCCTGTGCT CCATCTGAG CTCTGCACT TCGGCCATCC 780  
 65 AGAAGCTCC CACGGCGCCC CCGAGCGCC AGCCCCAGT GGACAGAAC AAGCTCTCCA 840  
 GGGCCACCA CGGGGAAGCA GGACATCACC TCCAGAGCC CAGCAGTAGG GAGACTGGCC 900  
 GAGGTGCCAA GGTGAGCGA GGTAGCAAGA GCCACCCAAA CGCCATGCC CGAGGAGAG 960  
 TCGGGGGCCT TGGGGCTCAG GGACCTTCG GAAGCAGCGA GTGGGAAGAC GAACAGTCTG 1020  
 AGTATTCTGA TATCGGAGG TGAATGAAA GGCCTGGCCA CGAAATCTTT CCTCCAGCC 1080  
 70 GTCCATTTC TATCTATGG ACATTCCAAA ACATTACCA TTAGAGAGGG GGGATGTAC 1140  
 ACGCAGGATT CTGTGGGAC TGTGGACTTC ATCAGGTGT GTGTTGCGG AACGACAGG 1200  
 TGAGATGGAG ACCCTCGGG CCGTGGGTC TCAGGGGTGC CTGTTGAATT CTGCATTAC 1260  
 ACGTACTCAA GGGAGCGCG CCGCGTTATC CTGTAACCT TGCTTCTTT CCATCTGTG 1320  
 AGTCAGTGGG TGTGGCGCG TCTGTTGTG GGGAGGTGAA CCAGGAGGG GCAGGGCAAG 1380  
 75 GCAGGGCCCC CAGAGCTGGG CCACACAGTG GGTGCTGGG CTGCCCCGA AGCTTCTGT 1440  
 GCAGCAGCCT CTGTGCTGT CTCCGCGAA GTCAGGGCG CTGATTCCA GGACAGGAGT 1500  
 GAATGTAAAA ATAAATATCG CTTAGAATGC AGGAGAAGG TGAGAGGAG GCAGGGGCG 1560  
 AGGGGTGCT TGTGTCGAAA CTGAATTC A GTTCTGTG TGCGGCTTG CGGTTCAGAG 1620  
 CTCTTGGCGA GGTGGGAGG AGGAGTGTCA TTTCTATGT TAATTCTGA GCCATTGTAC 1680  
 80 TGTCTGGCT GGGGGGACA CTGTCCAAG GAGTGGCCCC TATGAGTTA TATTTTAAC 1740  
 ACTGCTTCAA ATCTGATT TCACTTTTTT ATTTATCCAG TTATATCTAC ATATCTGTCA 1800  
 TCTAAATAAA TGGCTTCAA ACAAGCAAC TGGGTCAAT AAACAGCTC AAAGGGGTT 1860  
 TAAAAAATA AAAAAAGGCC CATCTTTGA GGTGATTTT TCTTTTTTT AAGTCTTAT 1920  
 TAAAAAGCTA TCAACAGCG ACATAGCCAT ACATCTGACT GCCTGACATG GACTCTGCC 1980  
 85 CACTTGGGG AAACCTTATA CCCAGAGGAA AATACACACC TGGGAGTAC ATTTGACAAA 2040  
 TTTCCTTAG GATTTCGTTA TCTCACCTG ACCCTCAGCC AAGATTGGTA AAGCTGCTC 2100  
 CTGGCGATTC CAGGAGACC AGCTGGAAC CTGGCTCTC CATGTGAGG GATGGGAAAG 2160  
 GAAGAGAGG AATGAAGACT ACTTAGTAAT TCCATCAGG AAATGCTGAC CTTTACATA 2220

WO 02/086443

PCT/US02/12476

AAATCAAGGA GACTGCTGAA AATCTCTAAG GGACAGGATT TTCCAGATCC TAATTGGA 2280  
TTTAGCAATA AGGAGAGGAG TCCAGGGGA CAAATAAAGG CAGAGAGAGA GAGAGAGAGA 2340  
GGGAGAGGAA GAAAGAGAG AGAGAAAAGA GCCTCGTGCC

Seq ID NO: 429 Protein sequence  
Protein Accession #: NP\_003705

1 11 21 31 41 51  
MCAERLQGF M TLALVLATFD PARGTDATNP PEGPQDRSSQ QKGRLSLQNT AEIQHCLVNA 60  
GDVGGCVFEC FENNSCBIRG LHGICMTFLH NAGKFDAGK SFIKDALKCK AHALRRERFGC 120  
ISRKCPAIRE MVSQLQRECY LKHDLCAAQ ENTRVIVEMI HFKDLLLHEP YVDLVNLLLT 180  
CGEEVKCAIT HSVQVQCEQN WSLCSILSF CTSAIQKPPT APPERQPVQD RTKLSRAHHG 240  
EAGHHLPEPS SRETGRGAGK ERGSKSHRNA HARGRVGLG AQGPSGSSEN EDEQSEYSDI 300  
RR

Seq ID NO: 430 DNA sequence  
Nucleic Acid Accession #: NM\_005940  
Coding sequence: 23..1489

1 11 21 31 41 51  
AAGCCACGCA GCCCGGGGCG GATGGCTCC GCGCGCTGG CTCGCGAGCG CGGCGCGCGG 60  
CGCCCTCCTG CCGCGGATGC TGCTGCTGCT GCTCCAGCGG CCGCGCTGC TGGCCCGGGC 120  
TCTGCGCGCG GAGCTCCACC ACCTCCATGC CGAGAGGAGG GGGCCACAGC CCTGGCATGC 180  
AGCCCTGCCC ATGAGCCCGG CACCTGCCCC TGCCACGCGG GAAGCCCCCG GGCCTGCCAG 240  
CAGCCTCAGG CCTCCCGCT GTGGCGTGCC CGACCCATCT GATGGGCTGA GTGCCCGCAA 300  
CGACACGAA AGGTTCTGTC TTCTGGGCGG CGCTGGGAG AAGACGGACC TCACCTACAG 360  
GATCCTTCGG TTCCCATGGC AGTTGGTGCA GGAGCAGGTG CGGCAGACGA TGGCAGAGGC 420  
CCTAAGGTA TGAGCGCATG TGACGCCACT CACCTTTACT GAGGTGCACG AGGCGCGTGC 480  
TGACATCATG ATGACCTTGC CCAGGTACTG GCATGGGGAC GACCTGCGGT TTGATGGGCC 540  
TGGGGCATC CTGGCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT 600  
CGACTATGAT GAGACCTGGA CTATCGGGGA TGACCAAGGC ACAGACCTGC TGCAGGTGGC 660  
AGCCCATGAA TTTGGCCAG TGCTGGGCTC GCAGCACACA ACAGCAGCCA AGGCCCTGAT 720  
TCCGCGCTTC TACACCTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGGCGT 780  
GTCAACACCTA TATGGCCAGC CCTGGCCAC TGTCACCTCC AGGACCCAG CCCTGGGCCC 840  
CCAGGCTGGG ATAGACACCA ATGAGATTGC ACCGCTGGAG CCAGACGCCC CGCCAGATGC 900  
CTGTGAGGCC TCCTTTGAGC CGGTCTCCAC CATCCGAGGC GAGCTCTTTT TCTTCAAAGC 960  
GGGCTTTGTG TGGCGCTCC GTGGGGGCA GCTGCAGCCC GGCTACCCAG CATTTGGCTC 1020  
TCGCCACTGG CAGGACTGCG CCAGCCCTGT GGAGCTGCC TTGAGGATG CCCAGGGCCA 1080  
CATTGTGTTT TCCAAAGTCT CTCAGTACTG GGTGTACGAC GGTGAAAGC CAGTCTGGG 1140  
CCCCGCACCC CTCACCGAGC TGGGCTCTGT GAGGTTCGCC GTCCATGCTG CCTTGGTCTG 1200  
GGGTCCCGAG AAGACAAGGA TCTACTTCTT CCGAGGAGG GACTACTGGC GTTTCACCC 1260  
CAGCACCCGG CGTGTAGACA GTCCCGTGCC CCGAGGGCC ACTGACTGGA GAGGGGTGCC 1320  
CTCTGAGATC GAGCTGCGCT TCCAGGATGC TGATGGCTAT GCCTACTTCC TGGCGGGCG 1380  
CCTCTACTGG AAGTTTGACC CTGTGAAGGT GAAGGCTCTG GAAGGCTTCC CCGTCTGT 1440  
GGGTCTGAC TTCTTTGCT GTGCCGAGCC TGCCAACT TCTCTGAC CATGGCTGG 1500  
ATGCCCTCAG GGGTCTGAC CCTTGCCAG CCACGAATAT CAGGCTAGAG ACCCATGGCC 1560  
ATCTTTGTGG CTGTGGGAC CAGGATGGG ACTGAGCCCA TGTCTCTGC AGGGGATGG 1620  
GGTGGGTAC AACCACATG ACACTGCGG GGAGGGCCAC GCAGGTCTGT GTCACCTGCC 1680  
AGGACTGTC TCAGACTGGG CAGGGAGGCT TTGGCATGAC TTAAGAGGAA GGGCAGTCTT 1740  
GGGACCCGCT ATGCAAGTCC TGGCAAACT GGTGCCCTG TCTCATCCCT GTCCCTCAGG 1800  
GTAGACCATG GGCAGGACTG GGGGAACCTG AGTGTCTTGT CTGTATCCCT GTTGTGAGGT 1860  
TCCTTCCAGG GGTGGGACT GAAGCAAGGG TGCTGGGGCC CATGGCCCT CAGCCCTGGC 1920  
TGAGCAACTG GGTGTAGGG CAGGGCCACT TCCTGAGTCT AGTGTCTTGT AGGTGCTTGC 1980  
ATCTGTCTGC CTCTGGCTG ACAATCCTGG AAATCTGTTT TCCAGAAATCC AGGCCAAAAA 2040  
GTTCACAGTC AAATGGGAGG GGTATTCTT CATGCAGAG ACCCCAGGCC CTGGAGGCTC 2100  
CAACATACCT CAATCTGTC CCAGGCCGGA TCCTCCTGAA GCCCTTTTGC CAGCATGCT 2160  
ATCCTCCAAA GCCATTGTAA ATGTGTGTAC AGTGTGTATA AACCTTCTTC TTCTTTTTC 2220  
TTTTTAACT GAGGATTGTC ATTAAACACA GTTGTTTTCT

Seq ID NO: 431 Protein sequence  
Protein Accession #: NP\_005931

1 11 21 31 41 51  
MAFAANLRS AARALLPPL LLLLQPPPL ARALPPDVH LHAERRGPQ WHAALPSSPA 60  
PAPATQEAPR PASSLRPPRC GVPDPSDGLS ARNRQKRFVL SGRWKEKTDL TYRILRFPWQ 120  
LVQEQVRQTM AEALKVWSDV TPLTFTEVHE GRADIMIDFA RYWHGDDLFP DGPGLLAHA 180  
FFPKTHREBD VHFYDETWT IGDDQGTDL QVAHEFGHV LGLQHTTAAK ALMSAFYTFR 240  
YPLSLSPDDC RGQVLYGQP WPTVTSRTPA LGPQAGIDTN EIAPLEPDAP PDACEASFDA 300  
VSTIRGELFF FKAGFVWRIR GGQLQPGYPA LASRHWQGLP SPVDAAPEDA QGHIWFFQGA 360  
QYVVYDGEKP VLGFPALTEL GLVRFPVHAA LVWGPERNKI YFFRGRDYWR FHPSTRRVDS 420  
PVPFRATDWR GVPSEIDAAF QDADGYAYFL RGRLYWKFDP VKVKALEGFP RLVGPDFFGC 480  
ABPANTFL

Seq ID NO: 432 DNA sequence  
Nucleic Acid Accession #: NM\_024022  
Coding sequence: 202..1563

1 11 21 31 41 51  
ACCGGGCACC GGACGGCTCG GGTACTTTCC TTCTTAATTA GGTGATGCCC GTGTGAGCCA 60  
GGAAAGGGCT GTTGTATTGG GAAGCCAGTA ACATGTGGC TCACTATCTC TTCCGTGGTG 120  
CCATCTACAT TTTTGGGAT CGGGAATTAT GAGGTAGAGG TGGAGGCGGA GCCGGATGTC 180  
AGAGTCCCTG AAATAGTCAC CATGGGGGAA AATGATCCGC CTGCTGTTGA AGCCCCCTTC 240  
TCATTCCGAT CGCTTTTGG CCTTGATGAT TTGAAAATAA GTCTGTTGC ACCAGATGCA 300

	GATGCTGTTG	CTGCACAGAT	CCTGTCACTG	CTGCCATTGA	AGTTTTTTCC	AATCATCGTC	360
	ATTGGGATCA	TTGCATTGAT	ATTAGCACTG	GCCATTGGTC	TGGGCATCCA	CTTCGACTGC	420
	TACGGAAAT	ACAGATGTGG	CTCATCCTTT	AAGTGTATCG	AGCTGATAGC	TCGATGTGAC	480
5	GGAGTCTCGG	ATTGCAAAAG	CGGGGAGGAC	GAGTACCGCT	GTGTCCGGGT	GGTGGTCAAG	540
	AATGCCGTGC	TCCAGGTGTT	CACAGCTGCT	TCGTGGAAGA	CCATGTGCTC	CGATGACTGG	600
	AAGGGTCACT	TGGCAATATG	TGCCGTGTGC	CAACTGGGTT	TCCCAAGCTA	TGTGAGTTCA	660
	GATAAECTCA	GAGTGAAGTC	GCTGGAGGGG	CAGTTCGGGG	AGGAGTTTGT	GTCCATCGAT	720
	CACCTCTTGC	CAGATGACAA	GGTGACTGCA	TTACACCACT	CAGTATATGT	GAGGGAGGGA	780
10	TGTGCTCTG	GCCACGTGGT	TACCTTGCAG	TGCACAGCCT	GTGGTCATAG	AAGGGGCTAC	840
	AGCTCAGCGA	TGCTGGGTGG	AAACATGTCC	TTGCTCTGCG	AGTGGCCCTG	GCAGGCCAGC	900
	CTTCAGTTCC	AGGGCTTACCA	CCTGTGCGGG	GGCTCTGTCA	TCACGCCCTT	GTGGATCATC	960
	ACTGCTGCAC	ACTGTGTTTA	TGACTTGTAC	CTCCCCAAGT	CATGGACCAT	CCAGGTGGGT	1020
	CTAGTTTCCC	TGTTGGACAA	TCCAGCCCCA	TCCCCTTGG	TGGAGAAGAT	TGTCTACCAC	1080
15	AGCAAGTACA	AGCCAAAGAG	GCTGGGCAAT	GACATCGCCC	TTATGAAGCT	GGCCGGGCCA	1140
	CTCAGTTTCA	ATGAAATGAT	CCAGCCTGTG	TGCCCTGCCA	ACTCTGAAGA	GAACCTCCCC	1200
	GATGGAAAAG	TGTGCTGGAC	GTGAGGATGG	GGGGCCACAG	AGGATGGAGG	TGACGCCCTCC	1260
	CCGTGCTGAA	ACCAAGCGCG	CGTCCCTTTG	ATTTCCAACA	AGATCTGCAA	CCACAGGGAG	1320
	GTGTACGGTG	GCAATCATCT	CCCTCCATG	CTCTGCGCGG	GCTACCTGAC	GGGTGGCGTG	1380
20	GACAGCTGCC	AGGGGAGCAG	CGGGGGGCCC	CTGGTGTGTC	AAGAGAGGAG	GCTGTGGAAG	1440
	TTAGTGGGAG	CGACCAAGCTT	TGGCATCGGC	TGCGCAGAGG	TGAACAAGCC	TGGGGTGTAC	1500
	ACCGGTGTCA	CTCCTTTCCT	GGACTGGATC	CACGAGCAGA	TGGAGAGAGA	CCTAAAAACC	1560
	TGAAGAGGAA	GGGGACAAGT	AGCCACCTGA	GTTCCTGAGG	TGATGAAGAC	AGCCCGATCC	1620
	TCCCTGGGAC	TTCCTGTGAT	GAACCTGCAC	ACGAGCAGAC	ACCCCTTGAG	CTCTGAGTTC	1680
25	CGGCACCACT	AGCAGGCCCC	AAGAGGCAC	CCTTCCATCT	GATTCCAGCA	CAACCTTCAA	1740
	GCTGCTTTTT	GTTTTTTGTG	TTTTTGAGGT	GGAGTCTCGC	TCTGTTGCC	AGGCTGGAGT	1800
	GCAGTGGCGA	AATCCCTGCT	CAGTGCAGCC	TCGCTTCCC	TGGTTCAAGC	GATTCTCTTG	1860
	CCTCAGCTTC	CCAGTAGCTG	GGGACCAAG	GTGCCGCCCA	CCACACCCAA	CTAATTTTTG	1920
	TATTTTTAGT	AGAGACAGGG	TTTACCATG	TTGGCCAGGC	TGCTCTCAAA	CCCTGACCTT	1980
30	CAATGATGAT	GCTGTCTTCA	GCCTCCACA	GTGCTGGGAT	TACAGGCATG	GGCCACCAAG	2040
	CCTAGCCTCA	CGCTCCTTTC	TGATCTTCAC	TAAGAACAAC	AGAAGCAGCA	ACTTGCAAGG	2100
	CGGGCCTTTC	CCACTGTGTC	ATCTGGTTTT	CTCTCCAGGG	GTCTTGCAAA	ATTCTGTAGC	2160
	AGATAAGCAG	TTATGTGACC	TCACGTGCAA	AGCCACCAAC	AGCCACTCAG	AAAAGACGCA	2220
	CCAGCCACGA	AGTGACAGAC	TGCAGTCACT	GCACGTTTTT	ATCTCTAGGG	ACCAGAACCA	2280
35	AACCCACCTC	TTCTACTTCC	AAGACTTATT	TTCAATGTGT	GGGAGGTTAA	TCTAGGAATG	2340
	ACTCGTTTAA	GGCCTATTTC	CATGATTTC	TGTAGCATT	TGGTGTCTGA	CGTATTATTG	2400
	TCCTTTGATT	CCAAATAATA	TGTTTCTTTC	CCTCAAAAAA	AAAAAATAAA	AAAAAATAAA	2460
	AAAAA						

Seq ID NO: 433 Protein sequence  
Protein Accession #: NP\_076927

	1	11	21	31	41	51	
45	MGENDPPAVE	APFSFRSLFG	LDDLKISPVA	PDADAVAAQI	LSLLPLKFFP	IIVIGIALI	60
	LALAIGLGIH	FDCSGKYRCR	SSFKEIELIA	RCDGVSDCKD	GEDEYRCVRV	GGONAVLQVF	120
	TAASWKTMCS	DDWKHYANV	ACAQLGFPSY	VSSDNLRVSS	LEGQFREEFV	SIDHLLPDDK	180
	VTALHHSVYV	REGCASHRVV	TLQCTACGHR	RGYSSRIVGG	NMSLLSQWFW	QASLOFQGYH	240
	LCCGSVITPL	WITTAHCVY	DLVLPKSWTI	QVGLVSLDND	PAPSHLVEKI	VYHSYKPKR	300
50	LGNDIALMLK	AGPLTFNEMI	QPVCLPNSSE	NFPDGKVCNT	SGMGATEDGG	DASPVNLHAA	360
	VPLISNKNICN	HRDVGGLIIS	PSMLCAGVLT	GGVDSQGDSD	GGPLVCQERR	LWKLVGATSF	420
	GIGCAEVNKP	GVTVTRVTSPL	DWIHBQMERD	LKT			

Seq ID NO: 434 DNA sequence  
Nucleic Acid Accession #: NM\_000493.2  
Coding sequence: 97..2139

	1	11	21	31	41	51	
60	CACCTTCTGC	ACTGCTCATC	TGGGCAGAGG	AAGCTTCAGA	AAGCTGCCAA	GGCACCATCT	60
	CCAGGAACCT	CCAGCACGCA	GAATCCATCT	GAGAATATGC	TGCCACAAAT	ACCCTTTTGT	120
	CTGCTAGTAT	CCTTGAACCT	GTTTCATGGA	GTGTTTTACG	CTGAACGATA	CCAAATGCCC	180
	ACAGGCATAA	AAGGCCCACT	ACCCAACACC	AAGACACAGT	TCTTCAATTCC	CTACACCATA	240
	AAGAGTAAAG	GTATAGCAGT	AAGAGGAGAG	CAAGGTACTC	CTGGTCCACC	AGGCCCTGCT	300
65	GGACCTGAG	GGCACCCAGG	TCCTTCTGGA	CCACCAGGAA	AACCAAGGCTA	CGGAAGTCTT	360
	GGACTCCAG	GAGAGCCAGG	GTTGCCAGGA	CCACCAGGAC	CATCAGCTGT	AGGGAACCA	420
	GGTGTGCCAG	GACTCCAGG	AAACCCAGGA	GAGAGAGGAC	CATATGGACC	AAAAGGAGAT	480
	GTGGACACAG	CTGGCCTACC	AGGACCCCGG	GGCCACCCAG	GACCACTTGG	AATCCCTGGA	540
	CCGGCTGGAA	TTTCTGTGCC	AGGAAAACCT	GGACAACAGG	GACCCACAGG	AGCCCCAGGA	600
	CCCAAGGGCT	TTCTGGGAGA	AAAGGGTGCA	CCAGGAGTCC	CTGGTATGAA	TGGACAGAAA	660
70	GGGGAAATGG	GATATGGTGC	TCCTGGTGGT	CCAGGTGAGA	GGGGTCTTCC	AGGCCCTCAG	720
	GGTCCACAG	GACCATCTGG	CCCTCCTGGA	GTGGGAAAAA	GAGGTGAAAA	TGGGGTTCCA	780
	GGACAGCCAG	GCATCAAGG	TGATAGAGGT	TTTCCGGGAG	AAATGGGACC	AATTGGCCCA	840
	CCAGGTCCCC	AAGGCCCTCC	TGGGGAACGA	GGGCCAGAAG	GCATTGGAAA	GCCAGGAGCT	900
	GCTGGAGCCC	CAGGCCAGCC	AGGGATTCCA	GGAAACAAAG	GTCTCCCTGG	GGCTCCAGGA	960
75	ATAGCTGGGC	CCCCAGGGCC	TCCTGGCTTT	GGGAACCCAG	GCTTGCCAGG	CCTGAAGGGA	1020
	GAAGAGGAGC	CTGCTGGCCT	TCCTGGGGGT	CCAGGTGCCA	AAGGGGAACA	AGGGCCAGCA	1080
	GGTCTTCTCG	GGAAAGCCAG	TCGACTGGA	CCCCCTGGGA	ATATGGGACC	CCAAGGACCA	1140
	AAAGGCATCC	CGGGTAGCCA	TGGTCTCCCA	GGCCCTAAAG	GTGAGACAGG	GCCAGCTGGG	1200
	CCTGCAGGAT	ACCTGGGGGC	TAAGGGTGAA	AGGGGTTCCC	CTGGGTGAGA	TGGAAAAACA	1260
80	GGGTACCCAG	GAACCCAGG	TCTGATGGT	CCTAAGGGTA	ACCCAGGGTT	ACCAGGTCCA	1320
	AAAGGTGATC	CTGGAGTTGG	AGGACCTCCT	GGTCTCCAG	GCCTGTGGG	CCCAGCAGGA	1380
	CCAAAGGGAA	TGCCCGGACA	CAATGGAGAG	GCTGGCCCAA	GAGGTGCCCC	TGGAAATACCA	1440
	GGTACTAGAG	GCCTATTGGG	GCCACAGGCT	ATTCCAGGAT	TCCTGGGGTC	TAAAGGGGAT	1500
85	CCAGGAAGTC	CCGGTCTCTC	TGGCCAGGCT	GGCATAGCAA	CTAAGGGCCT	CAATGGACCC	1560
	ACCGGGCCAC	CAGGGCTTCC	AGGTCCAAGA	GGCCACTCTG	GAGAGCCTGG	TCTTCCAGGG	1620
	CCCCCTGGGC	CTCCAGGGCC	ACCAGGTCAA	GCAGTCTATG	CTGAGGGTTT	TATAAGGCA	1680
	GGCCAAAGGC	CCAGTCTTTC	TGGGACCCCT	CTGTGTTAGTG	CCAACCCAGG	GGTAACAGGA	1740

ATGCTGTGT CTGCTTTTAC TGTATTCTC TCCAAAGCTT ACCCGCAAT AGGAACTCCC 1800  
ATACCATTG ATAAAATTG GTATAACAGG CAACAGCATT ATGACCCAG GACTGGAATC 1860  
TTTACTTGT AGATACCAGG AATATACTAT TTTTCATACC ACGTGCATGT GAAAGGGACT 1920  
CATGTTTGGG TAGGCTGTG TAAGAATGGC ACCCTGTAA TGTACACCTA TGATGAATAC 1980  
5 ACCAAAGCT ACCTGGATCA GGCTTCAGGG AGTGCCATCA TGATCTCAC AGAAAATGAC 2040  
CAGGTGTGGC TCCAGCTTCC CAATGCCGAG TCAAATGGCC TATACTCCTC TGATATGTC 2100  
CACTCCTCTT TCTCAGGATT CTTAGTGGCT CCAATGTGAG TACACCCAC AGAGCTAATC 2160  
TAAATCTTGT GCTAGAAAA GCATTCTCTA ACTCTACCCC ACCCTACAAA ATGCATATGG 2220  
10 AGGTAGGCTG AAAAGAATGT AATTTTATT TTCTGAAATA CAGATTGAG CTATCAGACC 2280  
AACAAACCTT CCCCCTGAAA AGTGAGCAGC AACGTAAAAA CGTATGTGAA GCCTCTCTTG 2340  
AATTTCTAGT TAGCAATCTT AAGGCTCTTT AAGGTTTTCT CCAATATTAA AAAATATCAC 2400  
CAAGAAGTCT CTGCTATTG TTTTCTAAC TATGAATATG AGAAGCAAA AAAAATAAT 2460  
TAAAAAATA AACAGAAATA GAGCTCTAAG TTATGTGAAA TTTGATTGTA GAAACTCGGC 2520  
15 ATTTCTTTT TAAAAAGCC TGTCTTCTAAC TATGAATATG AGAAGTCTA GAAACATCC 2580  
AGGAGTATC ATATACTTT GTAGAACTTA AATCTGTGAA TATTCAAATT TAAAGACAC 2640  
TGTATCCCTT AAAATATTTC TGATGGTGA CTACTCTGAG GCCTGTATGG CCCCCTTCAT 2700  
CAATATCTAT TGAATATAC AGGTGCATAT ATACTTGTTA AAGCTCTTAT ATAAAAAGC 2760  
CCCAAAATAT TGAAGTTCAT CTGAAATGCA AGGTGCTTTC ATCAATGAAC CTTTCAAAA 2820  
CTTTCTATG ATTGCAGAGA AGCTTTTAT ATACCCAGCA TAACTGGAA ACAGGTATCT 2880  
20 GACCTATCTT TATTTATGTA ACACAAAGT GATTAAATTG ATTTCTTTAA TTCCTTATG 2940  
AATCTTATGT GATATGATT TCTGGATTGA CAGAACATTA GCACATGTAC CTTGTGCCTC 3000  
CCATTCAAGT GAAGTATATA TTTCACTGA GGGTTTCAA ATTCGACTAG AAGTGGAGAT 3060  
ATATTATTTA TTTATGACT GTACTGTATT TTTATATTGC TGTTTAAAC TTTTAAGCTG 3120  
TGCTCACTT ATTAAGCAC AAAATGTTT ACCTACTCCT TATTTACGAC ACAATAAAT 3180  
25 AACATCAATA GATTTTAGG CTGAATTAAT TTGAAAGCAG CAATTGTCTG TTCTCAACCA 3240  
TTCTTCAAG GCTTTTCATT CGACAATAA AAATAACATC AATAG

Seq ID NO: 435 Protein sequence  
Protein Accession #: NP\_000484.2

1 11 21 31 41 51  
MLPQIPFLLL VSLMLVHGVF YAERYQMTG IKGPLPNTKT QFFIPTYIKS KGLAVRGEQG 60  
35 TPGPFGPAGP RGHGPGSPGL GKPGYGPGL QGEPGLPGPP GPSAVGKPGV PGLPGKPGER 120  
GPYGPFGDVG PAGLPGPRGP PGPPGIPGPA GISVPGKPGQ QGPTGAPGR GPFGEKAPG 180  
VPGMNGGKGE MGYGAPRGP ERGLPGPQGP TGPSGPPGVG KRGENGVPGP PGIKGDRGFP 240  
GEMGPFGPFG PQGPPGERGP BGIGKPGAAG APQGPPIGPT KGLPGAPGLA GPPGPPGPGK 300  
PGLPGLKGER GPAGLPGGPG AKGEQGPAGL PKPGLTGP GNMGPQGPGL IPGSHGLPGP 360  
40 KGETGPAGPA GYPGAKGERG SPGSDGKPGY PKPGLDGP KGNLPGPKG DPGVGGPPGL 420  
PGPVGPAGAK GMPGHNGEAG PRGAPGIPGT RGPFGPPGIP GFPGSKDGP SPGPPGPGI 480  
ATKGLNGPTG PPGPPGPRGH SGEPLPGPP GPPGPPQAV MPBGFIKAGQ RPSLSGTPLV 540  
SANGQVTGMP VSAFTVILSK AYPATGPIF FDKILYNRQQ HYDPRGTGIFT CQIPGIYYFS 600  
YHVHVKGTHV WVGLYKNGTP VNYTYDEYTK GYLDQASGSA IIDLTENDQV WLQLPNAESN 660  
45 GLYSSEYVHS SFGSGLVAPM

Seq ID NO: 436 DNA sequence  
Nucleic Acid Accession #: XM\_062811  
Coding sequence: 1..888

1 11 21 31 41 51  
ATGTGGGGCG CTGCGCGCTC GTCCGTCTCC TCATCCTGGA ACGCGGCTTC GCTCCTGCAG 60  
50 CTGCTGTCTGG CTGCGCTGCT GGCGGCGGGG GCGAGGGCCA GCGGCGAGTA CTGCCACGGC 120  
TGGCTGGAAG CCGAGGGCGT CTGGCGCATC GGCCTTCAGT GTCCCGAGCG CTTGACAGGC 180  
55 GGCGACGCCA CCATCTGCTG CCGCAGCTGC GCGTTGCGCT ACTGCTGCTC CAGCGCCGAG 240  
GCGCGCTCGG ACCAGGGCGG CTGCGACAAT GACCGCCAGC AGGCGGCTGG CGAGCTGGC 300  
CGGGGCGACA AAGACGGCCC CGACGGCTCG GCAGTGCCCA TCTACGTGCC GTTCTCTATT 360  
GTTGGCTCGG TGTTTGTGCG CTTTATCATC TTGGGGTCCC TGTGTGGCAGC CTGTGTCTGC 420  
60 AGATGTCTCC GGCCTAAGCA GGATCCCGAG CAGAGCCGAG CCCCAGGGGG TAACCGCTTG 480  
ATGGAGACCA TCCCATGAT CCCCAGTGCC AGCACTTCCC GGGGGTGGTC CTCACGCCAG 540  
TCCAGCACAG CTGCCAGTTC CAGCTCCAGC GCCAACTCAG GGGCGCGGCG GCCCCCAACA 600  
AGGTACACAG CCAACTGTGT CTGCGCGGAA GGGACCATGA ACAACGTGTA TGTCAACATG 660  
CCCAAGAAAT TCTCTGTGCT GAACTGTGAG CAGGCCACCC AGATTGTGCC ACATCAAGGG 720  
CAGTATCTGC ATCCCCATA CGTGGGGTAC ACGGTGCAGC ACGACTCTGT GCCCATGACA 780  
65 GCTGTGCCAC CTTTCTGGA CGGCGTGCAG CCTGGCTACA GGCAGATTCA GTCCCCCTTC 840  
CCTCACACCA ACAGTGAACA GAAGATGTAC CCAGCGGTGA CTGTATAA

Seq ID NO: 437 Protein sequence  
Protein Accession #: XP\_062811

1 11 21 31 41 51  
MWGARRSSVS SSNNAAALLQ LLLAALLAAG ARASGEYCHG WLDAQSVWRI GPQCPERFDG 60  
70 GDATICCGSC ALRYCCSAE ARLDQGGCDN DRQQGAGEPG RADKDGPDGS AVPIYVFFLI 120  
VGSVFVAFII LGLVAACCC RCLRPKQDPQ QSRAPGGNRL METIPMIPSA STSRGSSSRQ 180  
SSTAASSSSS ANGARAPPT RSQTNCLPE GTMMNVVNM PTNFVNLNQ QATQIVPHQG 240  
QYLEPPYVGY TVQHSVPMT AVPPFMDGLQ PGYRQIQSPF PHINSEKMY PAVTV

Seq ID NO: 438 DNA sequence  
Nucleic Acid Accession #: NM\_004004.1  
Coding sequence: 1..681

1 11 21 31 41 51  
85 ATGGATTGGG GCACGCTGCA GACGATCCTG GGGGGTGTGA ACAAACTC CACCAGCATT 60  
GGAAAGATCT GGCTCACGCT CCTCTTCATT TTTCGATTA TGATCCTGT TGTGGCTGCA 120  
AAGGAGGTGT GGGGAGATGA GCAGGCCGAC TTGTCTGCA ACACCTGCA GCCAGGCTGC 180

AAGAAGCTGT GCTACGATCA CTACTTCCCC ATCTCCCAAC TCCGGCTATG GGCCTGCGAG 240  
 CTGATCTTGG TGTCCAGCCC AGCGCTCCTA GTGGCCATGC ACCTGGCCTA CCGGAGACAT 300  
 GAGAAGAAGA GGAAGTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG 360  
 ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCCTGTGTG GGACCTACAC AAGCAGCATC 420  
 TTCTTCGGGG TCATCTTCCA AGCGGCTTTC ATGTACGTCT TCTATGTCTAT GTACGACGGC 480  
 TTCTCCATGC AGCGGCTGGT GAAGTGCAC GCCTGGCCTT GTCCCAACAC TGTGGACTGC 540  
 TTGTGTCCC GGGCCACCGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA 600  
 ATTTGCATCC TGCTGAATGT CACTGAATTG TGTATTTCG TAATTAGATA TTGTCTCTGG 660  
 AAGTCAAAAA AGCCAGTTTA A

Seq ID NO: 439 Protein sequence  
 Protein Accession #: NP\_003995.1

1 11 21 31 41 51  
 MDWGTQLTLL GGVNKHSTSI GKINLTVLFI FRIMILVVAA KEVWGDEQAD FVCNTLQPGC 60  
 KNVCYDHYFP ISHRLWALQ LIFVSSPALL VAMRVAYRRH EKKRKFIRGE IKSEPKDIEE 120  
 IKTKQVRIEG SLWNTYTSSI FFRVIFEAPF MYVFVVMYDG FSRQLRVKCN AMPCNPVTDC 180  
 FVSRPTEKTV FTVPMIAVSG ICILLNVTEL CYLLIRYCSG KSKKPV

Seq ID NO: 440 DNA sequence  
 Nucleic Acid Accession #: XM\_061091.1  
 Coding sequence: 1..2481

1 11 21 31 41 51  
 ATGCCAAATA CTTCAGGAAC AACCAGGATT GAAATTGGC TTCTCCAAGA GCCGCCGGG 60  
 CACGAGAGCG TGGTCGCGCG TCTCCTCCG GTGAGTCCCA GCCCGAGTT GGCCTGCGCG 120  
 CCGGGTACCG CGCAGTGCCG GGCTCGCGAT GACCGATTCA CGCTCCGAT GATTGGAGGT 180  
 CAGATGCATG GTGAGAAGGT AGATCTCTGG AGCCTTGGT TTCTTTGCTA TGAATTTTAA 240  
 GTTGGGAAGC CTCTTTTGA GGCAGAACAA GTCCATGTAA GCAAGAAAC CATCGGGAAG 300  
 ATTTCACTGC CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTCTGTGTA 360  
 GATGGGTCTA ACAGCGTCCG GAAAGGGAGC TTTGAAAGT CCAAGCACTT TGCCATCACA 420  
 GTCTGTGACG GTCTGACATC CAGCCCGAG AGGGTCAGAG TGGGAGCACT CCAAGTTCAGT 480  
 TCCACTCTCT ATCTGGAATT CCCCCTGGAT TCATTTTCAA CCCAACAGGA AGTGAAGGCA 540  
 AGAATCAAGA GGAATGTTTT CAAAGGAGGG CGCACGAGGA CGGAACCTGC TCTGAAATAC 600  
 CTCTGCACA GAGGGTTCGC TGGAGGCAGA AATGCTTCTG TGCCCCAGAT CCTCATCATC 660  
 GTCACTGATG GGAAGTCCCA GGGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGGAAAGG 720  
 GGTGTCACTG TGTGTGCTGT GGGGGTCAGG TTTCCAGGT GGGAGGAGCT GCATGCACCTG 780  
 GCCAGCGAGC CTAGAGGGCA GCACGTGCTG TTGGCTGAGC AGGTGGAGGA TGCCACCAAC 840  
 GGGCTCTTCA GCACCTCAG CAGCTCGGCC ATCTGTCTCA GCGCCACGCC AGCTGGGAGC 900  
 CCGAGCTTGA TCTTCATGGA GCGGTAAATG GGCATCTCTC TGATAGGCCCT CTGTGACTCG 960  
 CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAATGCCCTC 1020  
 TGCCCGCTCG CCTTTGGAGG GAGGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCCAG 1080  
 GTGAGCTCTC TCTTCTGCTT GGACAGCTCT GCGGGCACCA CTCTGGAGCG CTCTCTGCGG 1140  
 GCCAAAGTCT TCGTGAAGCG GTTGTGCGGG GCGGTGCTGA GCGAGGACTC TCGGGCCCGA 1200  
 GTGGGTGTGG CCACATACAG CAGGAGCTGT CTGTGTGGCG TGCCCTGTGG GGAATACAG 1260  
 GATGTGCTGT ACCTGTGCTG GAGCCTCGAT GGCATTCCCT TCGTGGTGG CCCCACCTCG 1320  
 ACGGGCAGTG CCTTGGCGCA GGGGGCAGAG CGTGGCTTGG GGAGCGCCAC CAGGACAGGC 1380  
 CAGGACCGGC CAGTAGAGT GTTGGTTTTG CTCAGTAGT CACACTCCGA GGATGAGGTT 1440  
 GCGGGCCAG CGGTCACTC AAGGGGCGGA GAGCTGCTCC TGCTGGGTGT AGGCACTGAG 1500  
 GCGGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCAA AGCATGTGAT GGTCTACTCG 1560  
 GATCCTCAGG ATCTGTCTCA CCAATCCCTC GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG 1620  
 CGGCCAGGGT GCGGACACA AGCCCTGGAC CTCGTCTTCA TGTGGACAC CTCTGCTCTA 1680  
 GTAGGGCCCG AGAATTTTGC TCAGATGCAG AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT 1740  
 GAGGTGAACC CTGACGTGAC ACAGGTGCGC CTGTGTGTGT ATGGCAGCCA GGTGACAGCT 1800  
 GCCTTGGGGC TGGACACCAA ACCCAACCGG GCTGCGATGC TGGGGGCCAT TAGCCAGGCC 1860  
 CCCTACTTAG GTGGGTGGG CTCAGCCGGC ACCGCCCTGC TGACATCTTA TGACAAAGTG 1920  
 ATGACCGTCC AGAGGGGTGC CCGGCTGGT GTCCCAAG CTGTGTGGT GCTCACAGGC 1980  
 GGGAGAGGGG CAGAGGATGC AGCCGTCTCT GCCCAGAAGC TGAGGAACAA TGGCATCTCT 2040  
 GTCTGTGTGC TGGGCGTGGG GCCTGTCTTA AGTGAGGGTC TGGGAGGCT TGCAGGTCCC 2100  
 CGGATTTCCC TGATCCAGT GGCAGCTTAC CCGACCTGC GGTACCAACA GGACGTGCTC 2160  
 ATTGATGGC TGTGTGAGA AGCCAAGCAG CCACTCAACC TCTGCAAAAC CAGCCCGTGC 2220  
 ATGAATGAGG CCAGCTGCGT CTGCAAGAT GGGAGCTACC GCTGCAAGTG TCGGATGGC 2280  
 TGGGAGGGCC CCCACTGCGA GAACCGTGAG TGGAGCTCTT GCTCTGTATG TGTGAGCCAG 2340  
 GGATGGATTG TTGAGACGCC CCTGAGGCAC ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT 2400  
 ACCCTCCCA CCAACTACAG AGAAGGCCTG GGCAGTAAA TGGTGCCTAC CTCTGGAAT 2460  
 GTCTGCGCCC CAGTCTCTTA G

Seq ID NO: 441 Protein sequence  
 Protein Accession #: XP\_061091.1

1 11 21 31 41 51  
 MPNTSGTTRI EIWLLQEPFG HRALVAALLP VSPSPALALA PGYPPVPAAD DRPTLFMIGG 60  
 QMHGKVDLW SLGVLCYEPF VGKPPPEANE VHSVETIGK ISAASKMMNC SAAVDIMPLL 120  
 DGSNSVKGGS PERSKHPAIT VCDGLDISPE RVRVGAFFQS STPHLEPFLD SPSTQGEVKA 180  
 RIKRMVFKGG RTETELALYI LLHRGLPGGR NASVPQILII VTDGKSQGDV ALPSKQLKER 240  
 GVTVPFVGVV PPRWEELHAL ASERPGQHV L LAEQVEDATN GLFSTLSSSA ICSSATPAGS 300  
 PELVFMERLM GISLIGPCDS QPCQNGGTCV PEGLDGYOCL CPLAFGGEAN CALKLSLECR 360  
 VDLLFLDSS AGTTLDGFLR AKVFKRFRVR AVLSEDSRAR VGVATYSREL LVAVPVGEYQ 420  
 DVPDLVWSLD GIPFRGGPTL TGSALRQAAB RGFGSATRTG QDRPRRVVVL LTESHSEDEV 480  
 AGPARHARAR ELLLLGVGSE AVRALEBEIT GSPKHVMVYS DPQDLFNQIP ELQKLCRSQ 540  
 RPGCRTOALD LVFMLDTSAS VGPFNAQMQ SFVRSALQF EVNPDVTQV LVVYGSQVQT 600  
 AFGLDTRKPR AAMLRATSOA PYLGGVGSAG TALLHIYDRV MTVQRGARPG VPKAVVVLTG 660  
 GRGAEDAAPV AQKLRRNGIS VLVVGVGPVL SEGLRLLAGP RDSLIHVAAY ADLRYHQDVL 720

IEWLCEBAKQ PVNLCKPSPC MNBGSCVLQN GSYRCKCRDG WEGPHCENRE WSSSCVCSVQ 780  
 GWILETPLRH MAPVQEGSSR TPPSNYREGL GTEMVPTFWN VCAPGP

Seq ID NO: 442 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..2424

```

1      11      21      31      41      51
10  ATGCCCCCTT TCCTGTGCTT GGAGGCCGTC TGTGTTTTC TGTITTCAG AGTGCCCCCA 60
    TCTCTCCCTC TCCAGGAAGT CCATGTAAGC AAAGAAACCA TCGGGAAGAT TTCAGCTGCC 120
    AGCAAAATGA TGTGTGCTC GGCTGCAGTG GACATCATGT TTCTGTTAGA TGGGTCTAAC 180
    AGCGTCGGGA AAGGGAGCTT TGAAAGGTCC AAGCACTTTG CCATCAGAGT CTGTGACGGT 240
    CTGGACATCA GCCCCGAGAG GGTGAGAGTG GGAGCATTCC AGTTCAGTTC CACTCCTCAT 300
15  CTGGAATTCC CCTTGGATTG ATTTTCAACC CAACAGGAAG TGAAGGCAAG AATCAGAGG 360
    ATGGTTTTCA AAGGAGGGCG CAGCGAGAGC GAACCTGTCT TGAATACCT TCTGCACAGA 420
    GGGTTGCGTG GAGGAGAGAA TGCTTCTGTG CCCCAGATCC TCATCATCGT CACTGATGGG 480
    AAGTCCGAGG GGGATGTGGC ACTGCCATCC AAGCAGCTGA AGGAAAGGGG TGTCACTGTG 540
    TTTGCTGTGG GGTGCTAGTT TCCCAGGTGG GAGGAGCTGC ATGCACTGGC CAGCGAGCCT 600
20  AGAGGGCAGC ACCTGCTGTT GGCTGAGCAG GTGAGGATG CCACCAACGG CCTCTTCAGC 660
    ACCCTCAGCA GCTCGGCCAT CTGCTCCAGC GCCACGCCAG ACTGCAGGGT CGAGGCTCAC 720
    CCTGTGAGC ACAGGAGCCT GGAGATGGTC CGGAGTTTCG CTGGCAATGC CCCATGCTGG 780
    AGAGGATGCG GGGGACCTCT TGGGTGCTG GCTGCACACT GTCCCTTCTA CAGCTGGAAG 840
    AGAGTGTTCG TAACCCACCC TGCCACCTGC TACAGGACCA CAGTCCGAGG CCCCCTGTAC 900
25  TCGCAGCCCT GCCAGAAATG AGGCACATGT GTTCCAGAA GACTGGACGG CTACCAATGC 960
    CTTCTGCCCG TGGCCTTTGG AGGGGAGGCT AACTGTGCCC TGAAGCTGAG CTGGAATGC 1020
    AGGCTCGAAC TCCTCTTCTT GCTGGACAGC TCTGCGGGCA CCACTCTGGA CGGCTTCTCT 1080
    CGGGCCAAAG TCTTCTGAA GCGGTTTGTG CGGGCCGTCG TGAGCGAGGA CTCTCGGGCC 1140
    CGAGTGGGTG TGCCACATA CAGCAGGGAG CTGCTGTGGG CGGTGCTGTG GGGGAGTAC 1200
30  CAGGATGTGC CTGACCTGGT CTGAGCCTTC GATGGCATTG CCTTCCGTGG TGGCCCCACC 1260
    CTGACGGGCA GTGCTTTGCG GCAGGCGGCA GAGCGTGGT TCGGGAGCGC CACCAGGACA 1320
    GGCCAGGACC GGCACGCTAG AGTGGTGGTT TTGCTCACTG AGTCACACTC CGAGGATGAG 1380
    GTTGGGGGCG CAGCGCGTCA CGCAAGGGCG CGAGAGCTGC TCTGCTGGG TGTAGGCACT 1440
    GAGGCGGTGC GGGCAGAGCT GGAGGAGATC ACAGGCAGCC CAAAGCATGT GATGCTTAC 1500
35  TCGGATCCTC AGGATCTGTT CAACCAATC CCTGAGCTGC AGGGGAAGCT GTGCAGCCGG 1560
    CAGCGGCCAG GGTGCCGAGC ACAAGCCCTG GACCTGCTCT TCATGTTGGA CACCTCTGCC 1620
    TCAGTAGGGC CGAGAAATTT TGCTCAGATG CAGAGCTTTG TGAGAAGCTG TGCCTCCAG 1680
    TTTGAGGTGA ACCCTGACGT GACACAGGTC GGCCTGTGGG TGTATGGCAG CCAGGTGCAG 1740
    ACTGCTTGG GGGTGCACAC CAAACCCACC CGGCTGCGGA TGTGCGGGG CATTAGCCAG 1800
40  GCGCCCTACC TAGTGGGTG GGGCTCAGCC GGCACCGCCC TGCTGCACAT CTATGACAAA 1860
    GTGATGACCG TCCAGAGGGG TGCCCGGCTT GGTGTCCCA AAGCTGTGGT GGTGCTCACA 1920
    GCGGGGAGAG GCGCAGAGGA TGCAGCCGTT CCTGCCAGA AGCTGAGGAA CAATGGCATC 1980
    TCTGTCTTGG TGTGGGGCGT GGGGCTGTCT CTAAGTGAGG GTCTGCGGAG GCTTGCAGGT 2040
    CCGCGGGATT CCTGATCCA CGTGGCAGCT TACGCGAGCC TGCGGTACCA CCAGGACGTG 2100
45  CTCATTGAGT GGTGTGTGG AGAAGCCAAG CAGCCAGTCA ACCTCTGCAA ACCAGCCCG 2160
    TGCAATGAAT GGGCAGCTGT GGTCTGCGAG AATGGGAGCT ACCGCTGCAA GTGTGCGGAT 2220
    GGCTGGGAGG GCGCCCACTG CAGAAACGCT GAGTGGAGCT CTTGCTCTGT ATGTGTGAGC 2280
    CAGGGATGGA TTCTTGAGAC GCGCCTCAGG CACATGGCTC CCGTGAGGGA GGGCAGCAGC 2340
    CGTACCCCTC CCAGCAACTA CAGAGAAGGC CTGGGCACTG AAATGGTGCC TACCTTCTGG 2400
50  AATGCTGTG CCCCAGGTCC TTAG
  
```

Seq ID NO: 443 Protein sequence  
 Protein Accession #: Eos sequence

```

55  1      11      21      31      41      51
    MPPFLLEAV CVFLPSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN 60
    SVGKGSFERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEFLDSFST QQEVKARIKR 120
60  MVPKGRRTET ELALKYLLHR GLPGGRNASV PQILIIIVTDG KSQGDVALPS KQLKRGVTV 180
    PAVGVPRPRM ERLHALASEP RGQHVLLAQ VEDATNGLPS TLSSAICSS ATPDCRVEAH 240
    PCEHRTLEMV REFAGNAPCW RGSRRTLAVL AAHCPPYSWK RVFLTHPATC YRTTCTPGPCD 300
    SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEK RVDLLPLDLS SAGTTLDGFL 360
    RAKVVRKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT 420
65  LTGSALRQAA ERPGSATRT QDRPRRVVV LLTSESEDE VAGPARHARA RELLLGVGS 480
    EAVRAELEBI TGSPKHMVY SDPQDLFNQI PELQGLCSR QRPGRTOAL DLVFLDTS 540
    SVGPENFAQM QSFVRSCALQ FEVNPDTVQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ 600
    APYLGGVGS A GTALLHIYDK VMTVQRGARP GVPKAVVLT GGRGARDAAV PAQLRNNIGI 660
    SVLVVVGVPV LSEGLRLLAG PRDSLHVA A YADLRYHQDV LIENLCGEAK QPVNLCKPSP 720
70  CMNBGSCVLQ NGSYRCKCRD WEGPHCENR EMSSCVCSV QGWILETPLR HMAPVQEGSS 780
    RTPPSNYREG LGTEMVPTFW NVCAPGP
  
```

Seq ID NO: 444 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 89..2356

```

75  1      11      21      31      41      51
    GCGCCCTGGC CCGAGCCGCG CCCGGGCTG TGAGTAGAGC CGCCCGGGCA CCGAGCGCTG 60
    GTGCCCGCTC TCCTTCGGTT ATATCAACAT GCGCCCTTTC CTGTGTCTGG AAGCGCTCTG 120
80  TGTITTCCTG TTTTCAGAGT TGCCCCATC TCTCCCTCTC CAGGAAGTCC ATGTAAGCAA 180
    AGAAACCATC GGAAGATTT CAGCTGCCAG CAAATGATG TGGTGCTCGG CTGCAGTGGA 240
    CATCATGTTT CTGTTAGATG GGTCTAACAG CTTGCGGAAA GGGAGCTTTG AAGGTCCAA 300
    GCACITTGCC ATCAGAGTCT GTGACGGTCT GGACATCAGC CCGGAGAGGG TCAGAGTGGG 360
85  AGCATTCAGC TTCAGTTCCA CTCCTCATCT GGAATTCCCC TTGATTTCAT TTTCAACCCA 420
    ACAGGAAGTG AAGGCAAGAA TCAAGAGGAT GGTITTCAAA GAGGAGGCGCA CGGAGACGGA 480
    ACTTGCTCTG AAATACCTTC TGACAGAGG GTTGCTGGA GGCAGAAATG CTTCTGTGCC 540
    CCAGATCCTC ATCATCGTCA CTGATGGGAA GTCCAGGGG GATGTGGCAC TGCCATCCAA 600
  
```

GCAGCTGAAG GAAAGGGGTG TCACTGTGTT TGCTGTGGGG GTCAGGTTTC CCAGGTGGGA 660  
 GGAGCTGCAT GCACCTGGCCA GCGAGCCTAG AGGGCAGCAC GTGCTGTGG CTGAGCAGGT 720  
 GGAGGATGCC ACCAACGGCC TCTTCAGCAC CCTCAGCAGC TOGGCCATCT GCTCCAGCGC 780  
 CACGCCAGAC TGCAGGTGTG AGGCTCACCC CTGTGAGCAC AGGACGCTGG AGATGGTCCG 840  
 GGAGTTGCTG GGCATGCCCC CATGCTGGAG AGGATCGCGG CGGACCCCTG CGGTGCTGGC 900  
 TGCACTCTGT CCCTTCTACA GCTGGAAGAG AGTGTTCCTA ACCCACCTTG CCACCTGCTA 960  
 CAGGACCAAC TGCCAGGGCC CCGTGACTC GCAGCCCTGC CAGAATGGAG GCACATGTGT 1020  
 TCCAGAAGGA CTGAGCGGCT ACCAGTCCCT CTGCCCCGCT GCCTTTGGAG GSGAGGCTAA 1080  
 CTGTGCCCTG AAGCTGAGCG TGAATGCAG GGTGACCTC CTCTTCTGCT TGGACAGCTC 1140  
 TGCGGGCACC ACTCTGAGCG GCTTCCTGCG GGCCAAAGTC TCGTGAAGC GGTTTGTGCG 1200  
 GGCGGTGCTG AGCGAGGACT CTCGGGCCCC AGTGGGTGTG GCCACATACA GCAGGGAGCT 1260  
 GCTGTGGGCG GTGCTGTGG GGGAGTACCA GGATGTGCCT GACCTGGTCT GGAGCCTCGA 1320  
 TGGCATTCCC TTCCGTGGTG GCCCCACCTT GACGGGCACT GCCTTGGGCG AGGCGGCAGA 1380  
 GCGTGGCTTC GGGAGGCGCA CCAGGACCGG CCAGGTAGAG TGGTGGTTTT 1440  
 GCTCACTGAG TCACATCCCG AGGATGAGGT TGCGGGCCCA GCGCGTCAGC CAAGGGCGCG 1500  
 AGAGCTGCTC CTGCTGGGCT TAGGCAGTGA GGCCGTGCGG CAGAGAGCTGG AGGAGATCAC 1560  
 AGGCAGCCCA AAGCATGTGA TGGTCTACT GGATCTCAG GATCTGTCA ACCAAATCCC 1620  
 TGAAGTCCAG GSGAAGCTGT GCAGCGGCA CCGGCCAGG TGCCGACAC AAGCCCTGGA 1680  
 CCTGCTCTTC ATGTTGGACA CCTCTGCCTC AGTAGGGCCC GAGAATTTTG CTCAGATGCA 1740  
 GAGCTTTGTG AGAAGCTGTG CCTTCCAGTT TGAGGTGAAC CCTGAGGTGA CACAGGTGG 1800  
 CCTGTGGTGT TATGGGAGCC AGGTGCAGAC TGCTTCCGGT CTGGACACCA AACCCACCGC 1860  
 GGCTGGGATG CTGCGGGCCA TTAGCCAGGC CCCCTACCTA GGTGGGGTGG GCTCAGCCGG 1920  
 CACGCGCTG CTGCACATCT ATGACAAAGT GATGACCGTC CAGAGGGGTG CCGGGCTGG 1980  
 TGTCCCAAAA GCTGTGGTGT TGCTCACAGG CCGGAGAGGC GCAGAGGATG CAGCCGTTC 2040  
 TGCCAGAGAG CTGAGGAACA ATGGCATCTC TGCTTGGTGT GTGGGGTGG GGCCTGTCT 2100  
 AAGTGAGGT CTGCGGAGCT TTGCAGGTCC CCGGATTCC CTGATCCAG TGGCAGCTTA 2160  
 CGCCGACCTG CGGTACCCAC AGGACGTGCT CATGAGTGG CTGTGTGGAG AAGCCAAAGCA 2220  
 GCCATCAAC CTCTGCAAC CCAGCCCGTG CATGAATGAG GGCAGCTGCG TCCTGCAGAA 2280  
 TGGGAGTAC CGCTGCAAGT GTCGGGATG CTGGGAGGGC CCCACTGCG AGAACCGATT 2340  
 CTGTGAGACG CCTTGAAGCA CATGGCTCCG GTGCAGGAG GCAGCAGCG TACCCTTCCC 2400  
 AGCAACTACA GAGAAGGCTT GGGCACTGAA ATGGTGCTTA CCTTCTGGAA TGTCTGTGCC 2460  
 CCAGGTCTCT AGAATGTCTG CTTCCGCGCG TGCCAGGAC CACTATTCTC ACTGAGGGAG 2520  
 GAGGATGTCC CACTGCAGC CATGCTGCTT AGAGACAAGA AAGCAGCTGA TGTACCCAC 2580  
 AAACGATGTT GTTGAAAGT TTTGATGTGT AAGTAAATAC CCACTTTCTG TACCTGTGT 2640  
 GCCTGTGTA GGTATGTCA TCTGCCACCT TTCCCTTGAG GATAAACAG GGTCTCTGAA 2700  
 GACTTAAAT TAGCGGCTG ACCTTCTCTT GCACACAATC AATGCTGCGC AGAATGTTGT 2760  
 TGACACAGTA ATGCCAGCA GAGGCCCTTA CTAGAGCATC CTTTGGACGG

Seq ID NO: 445 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MPPFLLEAV CVPLFSRVFP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMPLLDGSEN 60  
 SVGKGSFERS KHFAITVCDG LDISPERVRV GAPQFSSTPH LEPLDPSFT QQEVKARIKR 120  
 MVFKGGRTET ELALKYLLHR GLPGGRNASV PQILIIIVTDG KSQGDVALPS KQLKRGVTV 180  
 FAVGVRRFPW KELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH 240  
 PCEHRTLEWV REFAGNAPCW RGSRRTLAVL AAHCFFYSWK RVFLTHPATC YRTTCFPGCD 300  
 SQPCQNGTCT VPBGLDGVC LCPAFGGEA NCALKLSLEK RVDLLFLDLS SAGTTLDDGFL 360  
 RAKVPVKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT 420  
 LTGSALRQAA ERGFGSATRT GQDRPRRVVV LLTESHSEDE VAGPARHARA RELLLGVGS 480  
 EAVRALEBI TGSPKHVMVY SDPQDLFNQI PELQGLKCSR QRPGRCTQAL DLVFMLDTS 540  
 SVGPENPQPM QSPFVSCALQ FBNPDVTQV GLVVYGSQVQ TAFGLDTKPT RAAMLRASIS 600  
 APVLGGVGS A GTALLHIYDK VMTVQRGARP GVPKAVVVLV GGRGAEDAAV PAQKLNNNGI 660  
 SVLVGVGVFV LSEGLRLRLAG PRDSLHVAA YADLRVHODV LIENLGEAK QPVNLCKPSP 720  
 CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR FLRRP

Seq ID NO: 446 DNA sequence  
 Nucleic Acid Accession #: NM\_031942.1  
 Coding sequence: 145..1260

1 11 21 31 41 51  
 CCGAGCCCC GCCCTCCGG GCGCGGGTGG GCGCGCCAG CCTGCCAGCC GCGCTGCTGC 60  
 TGCTCTCTCT GCTGTGGGAC GCGTGACGCG GCGGCTGCTC CGCTCTCCCC GCTCCAAGCG 120  
 CGCATCTGGG CACCGGCCAC CAGCATGGAC GCTCGCGCGG TGCCGAGAA AGATCTCAGA 180  
 GTAAAGAAGA ACTTAAAGAA ATTACAGATAT GTGAAGTTGA TTTCCATGGA AACCTGTCA 240  
 TCTCTGATG ACAGTTGTGA CAGCTTTGCT TCTGATAATT TTGCAACAC GAGGCTGCAG 300  
 TCAGTTCCGG AAGCGCTAG GACCCGCGAG CAGTGCAGGC ACTCTGAGCC TCTCAGGGTG 360  
 CGGATGAAGT TTCCAGCGCG GAGTACCAGG GGAGCAACCA ACAAAGAGC AGAGTCCCGC 420  
 CAGCCCTCAG AGAATCTCTG GACTGATTCC AACTCCGATT CAGAAGATGA AAGTGGAAATG 480  
 AATTTTTTGG AGAAAAGGCG TTTAAATATA AAGCAAAACA AAGCAATGCT TGCAAAACTC 540  
 ATGTCTGAAT TAGAAAGCTT CCCTGGCTCG TTCCGTGGAA GACATCCCCC CCCAGGCTCC 600  
 GACTCACAAT CAAGGAGACC GCGAAGGCGT ACATTCCCGG GTGTGTCTTC CAGGAGAAAC 660  
 CCTGAACGGA GAGCTCGTCC TCTTACCAGG TCAAGGTCCC GGATCCTGG GTCCCTTGAC 720  
 GCTCTACCCA TGGAGGAGGA GGAGGAAGAG GATAAGTACA TGTGTGTGAG AAGAGGAGAG 780  
 ACCGTGGATG GCTACATGAA TGAAGATGAC CTGCCAGAA GCGTCTGCTC CAGATCATCC 840  
 TGACACCTTC CGCATATAAT TCGCCAGTGT GAAGAAATTA CAGAGGAGGA GTTGGAGAAC 900  
 GTCTGCAGCA ATTCTCGAGA GAAGATATAT AACCGTTAC TGGGCTCTAC TTGTATCAA 960  
 TGCCGTGAGA AGACTATTGA TACCAAAACA AACTGCAGAA ACCCAGACTG CTGGGGCGTT 1020  
 CGAGGCCAGT TCTGTGGCCC CTGCCCTCGA AACCGTTATG GTGAAGAGGT CAGGGATGCT 1080  
 CTGCTGGATC CGAATCTGCA TTGCCCGCCT TGTGAGGAA TCTGCACTG CAGTTTCTGC 1140  
 CGGCAGCGAG ATGGACGGTG TGGACTGGG GTCTTGTGT ATTTAGCCAA ATATCATGGC 1200  
 TTTGGGAATG TGATGCTCTA CTTGAAAAGC CTGAACAGG AATTTGAAAT GCAAGCATAA 1260  
 TATCTGAGAA TTCTGCTGCC TGCCCTTCTAC TTCTCAATC TTTCTGTGAA AAGTTTCCA 1320  
 TTTTCTACT GAAACCTCAG TTAATAATCT TGATGATCAG CTTGTTCAT AAGAACTCC 1380  
 AATCAAGITA ATCTTAGCAG ACATGTGTTT CTGGAGCATC ACAGAAGGTA TATTGCTAGT 1440



TACACTTTGC CCTCTGACG TTTCTTCTCT GCTCCCAACC CCCATCTCAT AGCATCCCCC 1500  
 TCTATTTCCA ATGCTCCTCT CCAACGCTT AGTTTCTGAA TTCTTTTAA ATTACAGTTT 1560  
 TATGAAAGCA TATTTTATTT ACTTGGTGTG GAAATAGCCC TCATAAAACC TAAGCACTTG 1620  
 GAAACACAAAT AATAGTATTA ACTAAGTGA TCTATTGAAT TTCAGAGAAG AGCCTTCTAA 1680  
 CTGTGTTTACA CAAAAACGAG TATGATTAG CACTCATACT AGTTGAAATT TTTAATAGAA 1740  
 TCAAGGCACA AAGTCTTAA AACCATGTGG AAAAATTAGG TAATTATTGC AGATTGATGT 1800  
 CTCTCAATCC CATGTATTGC GCATTATGTA CAAGTTGTTG TCACAGTTGA GACTTAATTT 1860  
 CTCTTAATTT CTCTGCCCG AAGGGTAAGT GGTGCGTCCA GCTTACAAGA TCATAATTC 1920  
 AAGGTTGGTG GGCAATGTAA TACTTAATTA AATAATGAT GGAAGAGCTA TCTGAGAGAT 1980  
 ATGAGTAAGC TGATTTGAAT TTTCAATATA AAACCTTAGT ATAATTGTAG TTTGCAAGT 2040  
 TTATTTTCACT TCACATGTAA GGTATTGCAA ATAAATCTT GGACAATTT GTATGGAAAC 2100  
 TTGATATTAA AAACAGTCT GTGTTCTTT GCAGTTTCTT GTAAATTTAT AAACAGGCA 2160  
 CAAGGTTCAA GTTTAGATT TAAGCACTTT TATAACAATG ATAAGTGCCT TTTTGAGAT 2220  
 GTAACTTTTA CAGATTGTTT AACCTGACAT CTCTGCCAGT CTAGTTTCTG GGCAGGTTTC 2280  
 CTGTGTCAGT ATTCCCTCTC CTCTTTGCAT TAATCAAGGT ATTGTGTAGA GGTGGAATCT 2340  
 AAGTGTGTTG ATGTCCAAAT TACTTGCATA TGTAAACCAT TGCTGTGCCA TTCAATGTTT 2400  
 GATGCATAAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTGTATCT GTAATGCTTT 2460  
 TATACAAAG TTTATTTTAA TAATAAAATG TTTGTCTTAA AAAAAAATAA

Seq ID NO: 447 Protein sequence  
 Protein Accession #: NP\_114148.1

1 11 21 31 41 51  
 MDARRVPQKD LRVKKNLKKF RYVKLISMET SSSSDSDSCDS FASDNFANTR LQSVREGCRT 60  
 RSQCRHSGPL RVAMKFPARS TRGATNKKAE SRQPSNSVT DSNDSSEDES GNMFLKRAL 120  
 NIKQNKAMLA KLMSSELESFP GSPFRHPLP GSDSQRRPR RRTFPGVASR RNPERRARPL 180  
 TRSRRIIGS LDALPMEEBE EEDKYLMLVRK RKTVDGYMNE DDLPRSRRSR SSVTLPHIIR 240  
 PVEEITEEL ENVCNSREK IYNSLGSST HQCRQKTYDT KTNCRNPDW GVRGQPCGFC 300  
 LRNRYGEEVR DALLDPNWHC PPCRGICNCS PCRQRDGRCA TGVLVYLAKY HGFNVHAYL 360  
 KSLKQEFEMQ A

Seq ID NO: 448 DNA sequence  
 Nucleic Acid Accession #: NM\_019894  
 Coding sequence: 1..1314

1 11 21 31 41 51  
 ATGTTACAGG ATCCTGACAG TGATCAACCT CTGAACAGCC TCGATGTCAA ACCCCTGCGC 60  
 AAACCCCGTA TCCCATGGA GACCTTCAGA AAGGTGGGGA TCCCATCAT CATAGCACTA 120  
 CTGAGCCTCG CGAGTATCAT CATTGTGTTT GTCCCTCATCA AGGTGATTCT GGATAAATAC 180  
 TACTTCTCTC CGCGGCGAGC TCTCCACTTC ATCCCGAGGA AGCAGCTGTG TGACGGAGAG 240  
 CTGGACTGTC CCTTGGGGGA GGACGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT 300  
 GCAGTGGCAG TCCGCTCTC CAAGSACCGA TCCACACTCG AGGTGCTGGA CTGGGCCACA 360  
 GGGAACTGGT TCTCTGCTG TTTGACAAC TTCACAGAAG CTCTGCTGA GACAGCCTGT 420  
 AGGCAGATGG GCTACAGCAG CAAACCCACT TTCAGAGCTG TGGAGATTGG CCCAGACCAG 480  
 GATCTGGATG TTGTGAAAT CACAGAAAAC AGCCAGGAGC TTGCGATCGG GAACCTCAAGT 540  
 GGGCCCTGTC TCTCAGGCTC CCTGGTCTCC CTGCACTGTC TTGCTGTGG GAAGAGCCTG 600  
 AAGACCCCCC GTGTGGTGGG TGGGGAGGAG GCCTCTGTGG ATTCTTGGCC TTGGCAGGTC 660  
 AGCATCCAGT ACGACAACA GCACGCTCTG GAGGGAGGCA TCCTGGACCC CCACTGGGTC 720  
 CTCACGGCAG CCCACTGCTT CAGGAAACAT ACCGATGTGT TCAACTGGAA GGTGCGGGCA 780  
 GGCTCAGACA AACTGGGCGA CTTCCCATCC CTGGCTGTGG CCAAGATCAT CATCATTGAA 840  
 TTCAACCCCA TGTACCCCAA AGACAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCACTC 900  
 ACTTTCTCAG GCACAGTCAG GCCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA 960  
 GCCACCCAC TCTGGATCAT TGGATGGGCG TTTACGAAGC AGAATGGAGG GAAGATGTCT 1020  
 GACATACTGC TGCAGGCGTC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGCAGACGAT 1080  
 GCGTACCAGG GGGAGTCCAC CGAGAAGATG ATGTGTGTCG GCATCCCGGA AGGGGGTGTG 1140  
 GACACCTGCC AGGGTGACAG TGGTGGGCCC CTGATGTACC AATCTGACCA GTGGCATGTG 1200  
 GTGGGCATCG TTAGCTGGGG CTATGGCTGC GGGGGCCCGA GCACCCACAG AGTATACACC 1260  
 AAGGTCTCAG CCTATCTCAA CTGGATCTAC AATGTCTGGA AGGCTGAGCT GTAA

Seq ID NO: 449 Protein sequence  
 Protein Accession #: NP\_063947.1

1 11 21 31 41 51  
 MLQDPDSQD LNSLDVKPLR KPRIPMETFR KVGPIIIIAL LSLASIIIV VLIKVILDKY 60  
 YFLCQPLHF IPRQLCDGE LDCPLGEDES HCVKSPFEGP AVAVRLSKDR STLQVLD SAT 120  
 GNVFSAFDN PTEALAEATC RQMGYSKPT FRAVEIGPDQ DLDVVEITEN SQELMRNNS 180  
 GPCLSGSLVS LHCLACGKSL KTRPVVGEE ASVDSWEPQV SIQYDKQHV GGSILDPHV 240  
 LTAHCFRKH TDVFNWVRA GSKLGSFSP LAVAKIIIE FNPMPKND IALMKLQFPL 300  
 TFSGTVRPIC LPFFDEELTP ATPLIWIGW FTXQNGGKMS DILLQASVQV IDSTRCNADD 360  
 AYQGEVTERK MCAGIEPEGV DTCQDSGGP LMYQSDQHWV VGIWSVGYC GGPSTPGVYT 420

Seq ID NO: 450 DNA sequence  
 Nucleic Acid Accession #: XM\_051860.2  
 Coding sequence: 52..3042

1 11 21 31 41 51  
 GCTCACCCAG GAAAAATATG CAATCGTCCC ATTGATATAC AGGCCACTAC AATGGATGGA 60  
 GTTAACCTCA GCACGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGCTAC 120  
 GACCGGGGCA GAGCTGCCG GAGCTACCGT GTACGGTTC TCTGTGGGAA GCCTGTGAGG 180  
 CCCAACTCA CAGTCAACAT GTGACCAAT GTGAACAGCA CCATTCTGAA CTGTGGAGAT 240  
 AATGTACAGT CATGGAACCC TGGAGATACC CTGGTCATTG CCAGTACTGA TTAATCCATG 300  
 TACCAGCAG AAGAGTTCCA GGTGCTTCCC TGCAGATCCT GCGCCCCCAA CCAGGTCAAA 360



	GTGGCAGGGA	AACCAATGTA	CCTGCACATC	GGGGAGGAGA	TAGACGGCGT	GGACATGCGG	420
	GCGGAGGTGG	GGCTTCTGAG	COGGAACATC	ATAGTGATGG	GGGAGATGGA	GGACAAATGC	480
	TACCCCTACA	GAACCCACAT	CTGCAATTTT	TTTGACTTCG	ATACCTTTGG	GGGCCACATC	540
5	AAGTTTGCTC	TGGGATTAA	GGCAGCACAC	TTGGAGGGCA	CGGAGCTGAA	GCATATGGGA	600
	CAGCAGCTGG	TGGGTCACTA	CCCGATTTCAC	TTCCACCTGG	COGCTGATGT	AGACGAAAGG	660
	GGAGGTTATG	ACCCACCCAC	ATACATCAGG	GACCTCTCCA	TCCATCATAC	ATTCTCTCGC	720
	TGCGTCAAG	TCCATGGCTC	CAATGGCTTG	TTGATCAAGG	ACGTTGTGGG	CTATAACTCT	780
	TTGGGCCACT	GCTTCTTCAC	GGAAAGATGG	CCGGAGGAAC	GCAACACTTT	TGACCACTGT	840
10	CTTGGCTCTC	TTGTCAAGTC	TGGAAACCTC	CTCCCTTCGG	ACCGTGACAG	CAAGATGTGC	900
	AAGATGATCA	CAGGAGACTC	CTACCCAGGG	TACATCCCCA	AGCCCAAGGA	AGACTGCAAT	960
	GCTGTGTCCA	CCTTCTGGAT	GGCCAATCCC	AACAACAACC	TCATCAACTG	TGCCGCTGCA	1020
	GGATCTGAGG	AAACTGGAT	TTGGTTTAT	TTTCAACAG	TACCAACGGG	CCCTCCGTG	1080
	GGAAATGTACT	CCCCAGTTA	TTGAGAGCAC	ATTCCACTGG	GAATAATTCTA	TAACAACCGA	1140
	GCACATTCCA	ACTACCGGGC	TGGCATGATC	ATAGACAACG	GAGTCAAAC	CACCGAGGCC	1200
15	TCTGCCAAGG	ACAAGCGGCC	GTTCCTCTCA	ATCATCTCTG	CCAGATACAG	CCCTCACCAG	1260
	GAGCGCGACC	CGCTGAAGAT	CCGGAGCGCG	GCCATCATCA	GACACTTCAT	TGCTTACAAG	1320
	AACCAAGGAC	ACGGGGCTTG	GCTGCGGGCG	GGGGATGTGT	GGCTGGACAG	CTGCGGTTT	1380
	GCTGACAATG	GCATTGGCCT	GACCTGGCC	AGTGGTGGAA	CCTTCCCGTA	TGACGACGGC	1440
20	TCCAAGCAAG	AGATAAGAA	CAGCTTGT	TTTGGCGAGA	GTGGCAACGT	GGGGACGGAA	1500
	ATGATGGACA	ATAGGATCTG	GGGCTCTGGC	GGCTTGGACC	ATAGCGGAAG	GACCTCCCT	1560
	ATAGGCCAGA	ATTTTCAAT	TAGAGGAATT	CAGTTATATG	ATGGCCCCAT	CAACATCCAA	1620
	AACCTGCACTT	TCCGAAGATT	TGTGGCCCTG	GAGGGCCGGC	ACACCAAGCG	CCTGGCCTTC	1680
	CGCTGAATA	ATGCTTGGCA	GAGCTGCCCC	CATAACAACG	TGACCGGCAT	TGCTTTGAG	1740
25	GACGTTCCGA	TTACTTCCAG	AGTGTCTCTC	GGAGAGCCTG	GGCCCTGGTT	CAACCAGCTG	1800
	GACATGGATG	GGGATAAGAC	ATCTGTGTTC	CATGACGTG	ACGGCTCCGT	GTCGAGTAC	1860
	CTGGCTCTCT	ACCTCACGAA	GAATGACAAC	TGGCTGGTCC	GGCACCAGCA	CTGCATCAAT	1920
	GTTCGCGACT	GGAGAGGGCG	CATTTCAGT	GGGTGCTATG	CACAGATGTA	CATTCAAGCC	1980
	TACAAGACCA	GTAACCTGGC	AATGAAGATC	ATCAAGAATG	ACTTCCCGAG	CCACCTCTT	2040
30	TACCTGGAGG	GGGCGCTCAC	CAGGAGCACC	CATTACCAGC	AATACCAACC	GGTTGTCAAC	2100
	CTGCAGAAAG	GCTACACCAT	CCACTGGGAC	CAGACGGCCC	CCGCCGAAC	CGCCATCTGG	2160
	CTCATCAACT	TCAACAAGAG	CGACTGGATC	CGAGTGGGCG	TGTGCTACCC	GCGAGGCAAC	2220
	ACATTCTCCA	TCTCTCGGGA	TGTTCACAAT	CGCCTGCTGA	AGCAAAAGTC	CAAGACGGGC	2280
	GCTTCTGTTA	GGACCTTGCA	GATGACAAA	GTGGAGCAGA	GCTACCTCTG	CAGGAGCCAC	2340
35	TACTACTGGG	ACGAGGACTC	AGGGCTGT	TTCTGGAAGC	TGAAGCTCA	GAACGAGAGA	2400
	GAGAAGTTTG	CTTTCTGTCT	CATGAAAGGC	TGTGAGAGGA	TAAAGATTAA	AGCTCTGATT	2460
	CCAAAGAACG	CAGGCTCAG	TGACTGCACA	GCCACAGCTT	ACCCCAAGTT	CACCGAGAGG	2520
	GCTGTGTTAG	ACGTGCCGAT	GGCCAAAGAG	CTCTTTGGTT	CTCAGCTGAA	AACAAGGAC	2580
	CATTCTCTGG	AGGTGAAGAT	GGAGAGTTCC	AAGCAGCACT	TCTTCCACCT	CTGGAACGAC	2640
40	TTGCTTACA	TTGAAGTGGA	TGGGAAGAG	TACCCAGTT	CGGAGGATGG	CATCCAGGTG	2700
	GTGGTGATTG	AOGGGAACCA	AGGGCGCGTG	GTGAGCCACA	CGAGCTTCAG	GAACCTCCAT	2760
	CTGCAAGGCA	TACCATGGCA	GCTTTTCAAC	TATGTGGCGA	CCATCCCTGA	CAATTCCATA	2820
	GTGCTTATGG	CATCAAGGGG	AAGATACGTC	TCCAGAGGCC	CATGAGCCAG	AGTGTGGA	2880
	AAGCTTGGGG	CAGACAGGGG	TCTCAAGTTG	AAAGAGCAAA	TGGCATTCGT	TGGCTTCAA	2940
45	GGCAGCTCC	GGCCCATCTG	GGTGACATG	GACACTGAGG	ATCACAAGC	CAAAATCTTC	3000
	CAAGTTGTGC	CCATCCCTGT	GGTGAAGAG	AAGAAATTGT	GAGGACAGCT	GGCGCCCGGT	3060
	GCCACCTCGT	GGTAGACTAT	GACGGTGACT	CTTGGCAGCA	GACCACTGGG	GGATGGCTGG	3120
	GTCCCCAGC	CCCTGCCAGC	AGCTGCTGGG	GAAGGCCGTT	TTTACGCCCT	GATGGGCCAA	3180
	GGGAAGGCTA	TACAGAGGCC	TGGTGCTGCC	ACCTGCCCTC	ACTCAAGTGT	CTACCTGAGC	3240
50	CCCTCGGGG	GGTGTGGCC	AATGTGGAA	ACATTCACTT	TCTCTCAGCC	TCTTGGGTGC	3300
	TTCTCTCTTA	TCTGTGCTC	TTCACTGGG	GTTTGGGGAC	CATATCAGGA	GACCTGGGTT	3360
	GTGCTGACAG	CAAGATCCA	CTTTGGCAGG	AGCCCTGACC	CAGCTAGGAG	GTAGTCTGGA	3420
	GGGCTGGTCA	TTCAAGATCT	CCCATGGTCT	TCAGCAGACA	AGTGAAGGTT	GTAATGTGAT	3480
	GAGAAAGAGC	CTTGGCCCTT	AGGAAATCTT	TACTCTCTGA	AGCAAGAGCC	AACTTCAAG	3540
55	GATTAGGAGC	TGGGGTAGAA	CTGGCTATCC	TTGGGGAAGA	GGCAAGCCCT	GGCTCTGGCC	3600
	GTGTCCACTT	TTCAAGGAC	TTTGAAGTGC	AGGTTTGAC	TTGACTAGA	TGACTCTCAA	3660
	AGGCCCTTTT	AGTTCTGAGA	TTCCAGAAAT	CTGCTGCATT	TCACATGGTA	CCTGGAAACC	3720
	AACAGTTTAT	GGATATCCAC	TGATATCCAT	GATGCTGGGT	GGCCAGCGC	ACACGGGATG	3780
	GAGAGGTGAG	AACATATGCC	TAGCTTGAGG	GGTCTGCACT	CCAGTAGGGC	AGGAGTCA	3840
60	GTCCATGTGC	ACTGCAATGC	CAGGTGGAGA	AATCAGAGAG	AGGTAATAATG	GAGGCCAGTG	3900
	CCATTTCAGA	GGGAGGGCTC	AGGAAGGCTT	CTTGTCTTCA	GGAATGAAGG	CTGGGGGCAT	3960
	TTTGTCTGGG	GGAGATGAGG	CAGCCTCTGG	AATGGCTCAG	GGATTAGGCC	CTCCCTGCGG	4020
	CTGCTGTCTG	AGCTGTGTGA	CTACGGGGTC	GGCCTTTGCT	CAGTCTCTCT	TGGCCCACTC	4080
	ATGATGGAGA	AGTGTGGTCA	GAGGGGAGCA	ATGGGCTTTG	CTGCTTATGA	GCACAGAGGA	4140
65	ATTCAAGTCCC	CAGGCAGCCC	TGCTCTGAC	TCCAAGAGGG	TGAAGTCCAC	AGAAGTGAAG	4200
	TCTTGCTTTA	GGGCTCAT	TGCTCTTCA	CCAGGGAAT	GAGCAGAGGG	GGCCTCCAGG	4260
	AGACCTTAGA	TGTGCTCGTA	CTCCCTCGGC	CTGGGATTTT	AGAGCTGGAA	ATATAGAAAA	4320
	TATCTAGCCC	AAAGCCTTCA	TTTTAACAGA	TGGGAAAGT	GAGCCCCCAA	GATGGGAAG	4380
	AACCAACAG	CTAAGGGAGG	GCCTGGGGAG	CCCCACCCTA	GGCCTTGTCT	CCACACCACA	4440
70	TTGCCCTAAC	AACCGGCCCC	AGAGTGCCCA	GGCACTCTCT	AGGTAGCTTC	TGGAAATGGG	4500
	GACAAGTCCC	CTGCAAGGAA	AGGAAATGAC	TAGAGTAGAA	TGACAGCTAG	CAGATCTCTT	4560
	CCCTCTGCT	CCACGGCAGC	ACAAACCCGC	CCTCCCTTTG	GTGTTGGCGG	TCCCTGTGGC	4620
	CTTCACTTTG	TTCACTACCT	GTGAGCCAG	CCTGGGTGCA	CAGTAGCTGC	AACTCCCAT	4680
	TGGTGTACC	TGGCTCTCTT	GTCTCTGAG	CTCTACAGGT	GAGGCCAGC	AGAGGGAGTA	4740
75	GGGCTCGCCA	TGTTTCTGGT	GAGCCAATTT	GGCTGATCTT	GGGTGTCTGA	ACAGCTATTG	4800
	GGTCCACCCC	AGTCCCTTTT	AGCTGCTGCT	TAATGCCCTG	CTCTCTCCCT	GGCCCACTT	4860
	ATAGAGAGCC	CAAGAGCTC	CTGTAAAGAG	GAGAACTCTA	TCTGTGGTTT	ATAATCTTGC	4920
	ACGAGGCACC	AGAGTCTCCC	TGGGTCTTGT	GATGAATAC	ATTTATCCCC	TTTCTCTGCC	4980
	CAACCACAAA	CTCTTCTCTT	CAAGAGGGC	CTGCTGGCT	CCCTCCACCC	AACTGCACCC	5040
80	ATGAGACTCG	GTCCCAAGAT	CCATTCCCCA	GGTGGGAGCC	AACTGTCAAG	GAGGTCTTTC	5100
	CCACCAACCA	CTTTTCAAGT	GCTGGGAGGT	GACCAATAGG	CTCTGCTTTT	AAAGATATGG	5160
	CTGCTTCAAA	GGCCAGAGTC	ACAGGAAGGA	CTTCTTCCAG	GGAGATTAGT	GGTGTGGAG	5220
	AGGAGAGTTA	AAATGACCTC	ATGTCCTTCT	TGTCACGGT	TTTGTGAGT	TTTCACTCTT	5280
	CTAATGCAAG	GGTCTCACAC	TGTGAACCA	TTAGGATGTG	ATCACTTTCA	GGTGGCCAGG	5340
85	AATGTTGAAT	GTCTTGTGCT	CAGTTCAATT	AAAAAGATA	TCTATTGAA	AGTTCTCAGA	5400
	GTGTGACATA	TGTTTACAG	TACAGGATCT	GTACATAAAA	GTTCCTTTCC	TAAACCATTC	5460
	ACCAAGAGCC	AATATCTAGG	CATTTTCTTG	GTAGCACAAA	TTTTCTTATT	GCTTAGAAAA	5520
	TGTGCTCCT	TGTTATTCTT	GTTTGTAA	CTTAAGTGAG	TTAGTCTTTT	AAGGAAAGCA	5580

WO 02/086443

PCT/US02/12476

ACGCTCCTCT GAAATGCTTG TCTTTTTTCT GTTGCCGAAA TAGCTGGTCC TTTTTCGGGA 5640  
 GTTAGATGTA TAGAGTGTGT GTATGTAAAC ATTTCTTGTA GGCATCACC A GAACAAAGA 5700  
 TATAATTTCT ATTTATTAT TATATGTGCA CTTCAAGAAG TCAGTGTGAG AGAAATAAAG 5760  
 AATTGTCTTA AATGTCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 451 Protein sequence  
 Protein Accession #: XP\_051860.2

1 11 21 31 41 51  
 MDGVNLSTEV VYKKGQDYRF ACYDRGRACR SYRVRFLECK PVRPKLTVTI DTNVNSTILN 60  
 LEDNVQSWKP GDTLVIASDT YSMYQAEFQ VLPSCRSCAFN QVKVAGKPMY LHIGEEIDGV 120  
 DMRAEVGLLS RNIIVMGEME DKCYPYRNHI CNFFDPTDFG GHIFALGFK AAHLEGTLEK 180  
 HMGOQVVGQY PIHFLHAGDV DERGGYDPT YIRDLSEIHT PSRCVTVHGS NGLLIKDVVG 240  
 YNSLGHCPFT EDGPEERNTF DHCLGLLVKS GTLLPSDRDS KMCKMITGDS YPGYIPKPRQ 300  
 DCNAVSTFWM ANPNMNLINC AAGSEETGF WFLFHHVPTG PSVGMYSPTY SEHPLGKFY 360  
 NNRAHSNYRA GMIIDNGVKT TEASAKDKRP FLSIIISARYS PHQDADPLKP REPAIRIRHI 420  
 AYKNQDHGAW LRGGDVWLDL CRPADNGIGL TLASGCTFFY DDGSKQEIKN SLFVGESGNV 480  
 GTEMMDNRIV GPGGLDHSGR TLPIGNFPI RGIQLYDGP I NIQNTCFRKP VALEGRHTSA 540  
 LAFRLNNAQ SCPHNNVTGI APEDVPITSR VFFGEPGPWF NQDMDGDKT SVPHVDGVS 600  
 SEYPGSYLTK NDNLVVRHFD CINVDPWRGA ICSGCYQMY IQAYKTSNLR MKIINKNDPFS 660  
 HPLVLEGALT RSTHYQYQYF VVTLQKGYTI HWDQTAPEL AIWLINFNKG DWIRVGLCYP 720  
 RGTTFSILSD VBNRLKQTS KTGVPVRTLQ MDKVEQSYFG RSHYYWDEDS GLFLKLLKQ 780  
 NEREKFAFCS MKGCERIKK ALIPKNAGVS DCTATAYPKP TERAUVVDVPM FKLLFGSQLK 840  
 TKDHFLEVKM ESSKQHFFHL WNDPAYIEVD GKYPSSDEG IQVVVDGNQ GRVVSHTSFR 900  
 NSILQGIWQ LFNYVATIPD NSIIVLMSKG RYVSRGPWTR VLEKLGADR G LKLEQMAFV 960  
 GFGKSRPIW VTLDTEHKA KIFQVVPFV VKKKKL

Seq ID NO: 452 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 261..2861

1 11 21 31 41 51  
 GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCCTGG CGAGCGCAAG 60  
 CGGCGCGGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCAGACT 120  
 AGCTACCACT CGCTTGCCCG ACGCCCGCGG AGCTGCGCGC GCGTGGCGGT CAGCGACCAAG 180  
 ACCTGCGGGG CGCTGCGGCT CCGTGGCGCG GAGGCGTGAC ACTGTCTCGG CTACAGACCC 240  
 AGAGGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCTC TTCAAGGCCA 300  
 TGCTGACCAT CAGCTGGCTC ACTCTGACCT GCTTCCTGGG GGCCACATCC ACAGTGGCTG 360  
 CTGGGTGCCC TGACCAAGAG CCTGAGTTGC AACCTTGGAA CCCTGGCCAT GACCAAGACC 420  
 ACCATGTGCA TATCGGCCAG GGCAAGACAC TGCTGCTCAC CTCTCTTGCC ACGGTCTATT 480  
 CCATCCACAT CTGAGGGGGA GGCAAGCTGG TCATTAAAGA CCACGACGAG CCGATTGTTT 540  
 TGCGAACCCG GCACATCCTG ATTGACAACG GAGGAGAGCT GCATGCTGGG AGTGCCTCT 600  
 GCGCTTTCCA GGGCAATTTC ACCATCATT TGTATGGAAG GGCTGATGAA GGTATTGAGC 660  
 CGATCTCTTA CTATGCTCTG AAGTACATTG GGGTGGTAA AGGAGGCGCT CTGAGTTGC 720  
 ATGGACAGAA AAGGCTCTCC TGGACATTTC TGAACAAGAC CCTTACCCA GGTGGCATGG 780  
 CAGAAGGAGG CTATTTTCTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA 840  
 TCGACCCCAA ATCAGGCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAAGA 900  
 AAGAGAGTGA ACGTCTGGTC CAGTATTGGA ACGCGTGCC CGATGGCAGG ATCCTTCTG 960  
 TTGCACTGAA TGATGAAGGT TCTGGAATTC TGGATGACAT GGCCAGGAAG GCGATGACCA 1020  
 AATTGGGAAG CAACACTTTC CTGACCTTGG GATTTAGACA CCCTTGGAGT TTCTTAACCT 1080  
 TGAAAGGAAA TCCATCATCT TCACTGGAAG ACCATATTGA ATATCATGGA CATCGAGGCT 1140  
 CTGCTGCTGC CGGTGTACGG TCTCTCTGTG GGAAGCCTGT GAGGCCCAA CTCACAGTCA 1200  
 CTTTGTCCAG TGAGTGGGTT CAAGACGTGG AGTGGACGGA GTGGTTCGAT CATGATAAAG 1260  
 TATCTCAGAG TAAAGGTGGG GAGAAAATTT CAGACCTCTG GAAAGCTCAC CCAGGAAAAA 1320  
 TATGCAATCG TCCATTTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGACCG 1380  
 AGGTGTCTA CAAAAAGGCG CAGGATTATA GGTTTGCTTG CTACGACCGG GGCAGAGCCT 1440  
 GCGGAGCTA CGGTGTACGG TTCTCTGTG GGAAGCCTGT GAGGCCCAA CTCACAGTCA 1500  
 CCATTGACAC CAATGTGAAC AGCACCATTG TGAACCTTGA GGATAATGTA CAGTATGGA 1560  
 AACCTGGAGA TACCCTGGTC ATTGCCAGTA CTGATTACTC CATGTACCA GCGAAGAGT 1620  
 TCCAGGTGCT TCCCTGCAGA TCCCTGCGCC CCAACCAAGT CAAAGTGGCA GGGAAACCA 1680  
 TGTAACCTGA CATCGGGGAG GAGATAGACG CGGTGGACAT CGCGGGCGAG GTTGGGCTTC 1740  
 TGAGCGGAA CATCATAGT ATGGGGGAGA TGGAGGACAA ATGCTACCCC TACAGAAACC 1800  
 ACATCTGCAA TTTCTTTGAC TTGATACCT TTGGGGGCA CATCAAGTTT GCTCTGGGAT 1860  
 TTAAGGCAGC ACATCTGGAG GGCACGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC 1920  
 AGTACCGAT TCACTTCCAC CTGGCCGCTG ATGTAGACGA AAGGGGAGGT TATGACCCAC 1980  
 CCACATACAT CAGGACCTTC TCCATCCATC ATACATTCTC TCGCTGCGTC ACAGTCCATG 2040  
 GCTCCATGG CTGTGTGATC AAGGACGTTG TGGGCTATAA CTCTTTGGGC CACTGCTTCT 2100  
 TCACGGAAGA TGGCCGGGAG GAACGCAACA CTTTTGACCA CTGCTCTGGC CTCCTGTGTA 2160  
 AGTCTGAAC CCTCTTCCCC TCGGACCGTG ACAGCAAGAT GTGCAAGATG ATCACAGAGG 2220  
 ACTCTTACCC AGGTATACAT CCCAAGCCCA GGCAGACTG CAATGCTGTG TCCACCTTCT 2280  
 GGATGGCCAA TCCCAACAAC AACCTCATCA ACTGTGCGCG TGCAGGATCT GAGGAAACTG 2340  
 GATTTTGGTT TATTTTTCAC CACGTACCAA CGGGCCCTTC CGTGGGAATG TACTCCCGAG 2400  
 GTTATTGAGA GCACATTCCA CTGGGAAAA TCTATAACAA CCGAGCACAT TCCAACTACC 2460  
 GGGCTGGCAT GATCATAGAC AACGGAGTCA AAACCAACGA GGCCTCTGCC AAGGACAAGC 2520  
 GGCCTGTCT CTCAATCATC TCTGCCAGAT ACAGCCCTCA CCAGGACGCC GACCCGCTGA 2580  
 AGCCCGGGA GCGGGCCATC ATCAGACACT TCATTGCCCTA CAGAACCAG GACCACGGGG 2640  
 CCTGGCTGCG CGCGGGGAT GTGTGGCTGG ACAGCTGCCA TTTCAGAGGG GAGGCTCAGG 2700  
 AAGGCTTCTT GCTTACAGGA ATGAAGGCTG GGGGCAATTT GCTGGGGGGA GATGAGGCAG 2760  
 CCTCTGGAAT GGCTCAGGGA TTCAGCCCTC CTGCGCGCTG CCGTCTGAAG CTGGTGAATA 2820  
 CGGGGTGCGC CTTTGTCTAC GTCTCTCTGG CCCACTCATG ATGGAGAAGT GTGGTCAGAG 2880  
 GGGAGCAATG GGCTTTGCTG CTTATGAGCA CAGAGGAATT CAGTCCCCAG GCAGCCCTGC 2940  
 CTCTGACTCC AAGAGGCTGA AGTCCACAGA AGTGAGCTCC TGCCTTAGGG CCTCATTTGC 3000  
 TCTTCATCCA GGAACCTGAG CACAGGGGCG CTCAGGAGA CCCTAGATGT GCTGTACTC 3060  
 CCTCGGCTG GGAATTCAGA GCTGGAAATA TAGAAAAATAT CTAGCCCAA GCCTTCATT 3120

TAACAGATGG GGAAGTGTAG CCCCCAAGAT GGGAAAGAAC CACACAGCTA AGGGAGGGCC 3180  
 TGGGGAGCCC CACCCTAGCC CTGCTGCTCA CACCACATTC CCTCAACAC CGGCCCCAGA 3240  
 GTGCCAGGCG ACTCCTGAGG TAGCTTCTGG AAATGGGGAC AAGTCCCTTC GAAGGAAAGG 3300  
 AAATGACTAG AGTAGAATGA CAGCTAGCAG ATCTCTTCCC TCCTGCTCCC AGCGCACACA 3360  
 AACCOCGCCCT CCCCCTGGTG TTGGCGGTCC CTGTGGCCTT CACTTTGTTT ACTACCTGTC 3420  
 AGCCAGGCTT GGGTGCACAG TAGCTGCAAC TCCCATTTGG TGCTACCTGG CTCTCCTGTC 3480  
 TCTGAGCTCT TACAGGTAGG GCCCAGCAGA GGGAGTAGGG CTGCCATGTT TTCTGGTGGG 3540  
 CCAATTGGCG TGATCTTGGG TGTCTGAACA GCTATTGGGT CCACCCCACT CCCTTTCAGC 3600  
 TGCTGCTTAA TGCCCTGTCT TCTCCCTGGC CCACCTTATA GAGAGCCCAA AGAGCTCCTG 3660  
 TAAGAGGGAG AACTCTATCT GTGGTTTATA ATCTTGCAAG AGGCACCAGA GTCTCCCTGG 3720  
 GTCTTGTGAT GAATACATTT TATCCCCCTT CTGCCCCCAA CCACAACTCT TTTCTTCAA 3780  
 AGAGGGCCTG CCTGGCTCCC TCCACCCCAAC TGCAACCCAT AGACTCGGTC CAAGAGTCCA 3840  
 TTCCCCAGGT GGGAGCCAAAC TGTCAAGGAG GTCTTTCCCA CCAAACTCTT TTCAGCTGCT 3900  
 GGGAGGTGAC CATAGGGCTC TGTCTTTTAA GATATGGCTG CTTCAAAGGC CAGAGTCACA 3960  
 GGAAGGACTT CTTCAGGGA GATTAGTGGT GATGGAGAG AGAGTTAAAT TGACCTCATG 4020  
 TCCTTCTGTT CCACGGTTTT GTTGAGTTTT CACTCTTCTA ATGCAAGGGT CTCACACTGT 4080  
 GAACCACTTA GGATGTGATC ACTTTCAGGT GGGCAGGAAT GTTGAATGTC TTTGGCTCAG 4140  
 TTCAATTAA AAAGATATCT AITTTAAAGT TCTCAGAGTT GTACATATGT TTCACAGTAC 4200  
 AGGATCTGTA CATAAAAGTT TCTTCTCTAA ACCATTCAAC AAGAGCCAAT ATCTAGGCAT 4260  
 TTTCTTGGTA GCACAAATTT TCTTATTGCT TAGAAAATTT TCCTCTCTGT TATTCTGTT 4320  
 TGTAAGACTT AAGTGAAGTT GGTCTTTAAG GAAAGCAAGC CTCCTCTGAA ATGCTTGTCT 4380  
 TTTTCTGTT TCCGAAATAG CTGCTCTTTT TTGGGAGGTT AGATGTATAG AGTGTTTGTA 4440  
 TGTAACACTT TCTTGTAGCG ATCAACATGA ACAAAGATAT ATTTCTATT TATTATTAT 4500  
 ATGTGCACTT CAAGAAAGTCA CTGTCAAGAG AATAAAGAT TGTCTTAAAT GTCATGATTG 4560  
 GAGATGCTCT TTGCAATTGCT TGGAGGGGGT GTACCTAGAG CCAAGGAAAT TGGCTCTGGT 4620  
 TTGGAAAAAT TTTGCTGTTA TTATAGTAAA CATACAAAGG ATGTCAAAAA AAAAAAATAA 4680  
 AAAAAAATAA AAAAAAATAA AA

Seq ID NO: 453 Protein sequence  
 Protein Accession #: Bos sequence

1 11 21 31 41 51  
 | | | | |  
 MGAAGRQDFL FKAMLTISWL TLTCFPGATS TVAAGCPDQS PELQPNPGE DQDHEVHIGQ 60  
 35 GTKILLTSSA TVVSIHISEG GKLVIKDDE PIVLRTRHIL IDNGGELHAG SALCPFGQNF 120  
 TIILYGRADE GIQDPYYGL KYIGVGKGA LELHGQKRLS WFLNKTLHP GGMABGGYFF 180  
 ERSWGHGVI VHVDPKSGT VIHSDRFDY RSKKESERLV QYLNVPDGR ILSVAVNDEG 240  
 SRNLDMARK AMTKLGSKEP LHLGPRHPWS FLTVMGNPSS SVEDHIEYHG HRGSAARVP 300  
 40 KLPQTEHGY FNVSLSEWV QDVETWTFD HDKVSQTKGG EKISDLWKAH PGKICNRPID 360  
 IQATTMDGVN LSTVTVYKKG QDYRFACVDR GRACRSYVR FLGCPVPRPK LTVTIDTNNV 420  
 STILNLEDNV QSWKPGDTLV IASTDYSMYQ AEEFQVLPGR SCAPNQVKA GKMPLYLHGE 480  
 EIDGDMRAE VGLLSRNIIV MGEMEDKCYR YRNHICNFFD FDTFPGHIFK ALGPKAHLR 540  
 GTELKMGQQ LVQYPIHFH LAGDVDERGG YDPPTYIRDL SIHHTFSRCV TVHGSNGLLI 600  
 45 KDVVYNSLG HCFPTEDGPE ERNTPDHCLG LLVKSQTLPL SDRDSKMKRM ITEDSPGYI 660  
 PKPRQCAV NLINCAAGS EETGFWFIFR HVPTGFSVGM YSPGYSEHIP 720  
 LGKFYNNRAH SNYRAGMIID NGVKTTEASA KDKRPFSLII SARYSPHQDA DPLKPREPAI 780  
 IRHFYAYKQV DHGAWLRGGD VWLDSCHFRG EAQEGFLITG MKAGGILLGG DEASAGMAQG 840  
 FSPPCRLKLV LVTGSPFAH VSLAHS

Seq ID NO: 454 DNA sequence  
 Nucleic Acid Accession #: NM\_013282.2  
 Coding sequence: 85..2466

1 11 21 31 41 51  
 | | | | |  
 CGACTCCTTA GAGCATGGCA TGGCTCAGAG GTGCTGGTAA AACTGATGGG GGTTTTGTCT 60  
 GTCCCTCCCC TCAGCGCCCA CACCATGTGG ATCCAGGTTT GGACCATGGA CGGGAGGCAG 120  
 ACCCACACGG TCGACTCGCT GTCCAGGCTG ACCAAGGTGG AGGAGCTGAG CGCGAAGATC 180  
 60 CAGGAGCTGT TCCAGCTGGA GCCAGGCTTG CAGAGGCTGT TCTACAGGGG CAAACAGATG 240  
 GAGGAGCGGC ATACCTCTTT CGACTACGAG GTCCGCTGTA ATGACACCAT CCAGCTCCTG 300  
 GTCCGCCAGA GCCTCGTGCT CCCCACAGC ACCAAGGAGC GGGACTCCGA GCTCTCCGAC 360  
 ACCGACTCGT GCTGCTGCTT GGGCCAGAGT GAGTCAGACA AGTCCTCCAC CCACGGCGAG 420  
 CGGCGCGCCG AGACTGACAG CAGGCCAGCC GATGAGGACA TGTGGGATGA GACGGAATTG 480  
 65 GGGCTGTACA AGGTCAATGA GTACGTGAT GCTGGGACA CGAACATGGG GGGCTGGTTT 540  
 GAGGCGCAGG TGGTCAGGCT GACGCGAAG GCCCCTCCC GGGACGAGCC CTCAGCTCC 600  
 ACGTCCAGCG CGGCGCTGGA GGAGGAGCTC ATTTACACG TGAATACGA CGACTACCG 660  
 GAGAAGCGCG TGGTCCAGAT GAACTCCAG GACGTCCGAG CGCGCGCCG CACCATCATC 720  
 AAGTGGCAGG ACCTGGAGGT GGGCCAGGTG GTCATGCTCA ACTACAACCC CGACAACCCC 780  
 70 AAGGAGCGGG GCTTCTGGTA CGACGCGGAG ATCTCCAGGA AGCGCGAGAC CAGGACGGCG 840  
 CGGGAACCTT ACGCCAACGT GGTGCTGGGG GATGATTCTC TGAACGACTG TCGATCATC 900  
 TTCTGGAGCG AAGTCTTCAA GATTGAGCGG CCGGGTGAAG GGAGCCCAT GGTTGACAAC 960  
 CCAATGAGAC GGAAGAGCGG GCCGTCTGTC AAGCACTGCA AGGACGACGT GAACAGACTC 1020  
 TGCCGGGTCT GCGCTGCCA CTTGTGGGGG GCGCGGAGG ACCCGACAA GCAGCTCATG 1080  
 75 TGCGATGAGT GCGACATGGC CTTCACATC TACTGCTTGG ACCCGCCCTT CAGCAGTGT 1140  
 CCCAGCGAGG ACGAGTGGTA CTGCCCTGAG TGCCGGAATG ATGCCAGGA GGTGGTACTG 1200  
 CGCGGAGAGC GGTGAGAGA GAGCAAGAAG AAGGCGAAGA TGCCCTCGGC CACATCGTCC 1260  
 TCACAGCGGG ACTGGGCAAA GGGCATGGCC TGTGTGGGCC GCACCAAGGA ATGTACCATC 1320  
 GTCCCGTCCA ACCACTACGG ACCATCCCG GGGATCCCCG TGGGCACCAT GTGGCGTTC 1380  
 80 CGAGTCCAGG TCAGCGAGTC GGGTGTCCAT CGGCCCCACG TGGCTGGCAT ACACGGCGGG 1440  
 AGCAACGAGG GAGCGTACTC CTTAGTCTTG CGGGGGGGCT ATGAGGATGA CGTGGACCAT 1500  
 GGGAAATTTT TCACATACAC GGTAGTGGT GGTGAGATCT TTTCCGCAAA CAAGAGGACC 1560  
 CGGGAACAGT CTTGTATCA GAAACTCACC AACACCAACA GGGCGCTGGC TCTCACTGC 1620  
 85 TTTGCTCCCA TCAATGACCA AGAAGGGGCC GAGGCCAAGG ACTGGCGGTC GGGGAAGCCG 1680  
 GTCAGGGTGG TGCCCAATGT CAAGGTGGC AAGAATAGCA AGTACGCCCC CGCTGAGGGC 1740  
 AACCGCTACG ATGGCATCTA CAAGGTGGT AAATACTGGC CCGAGAAGGG GAAGTCCGGG 1800  
 TTTCTGCTGT GGGCTACCT TCTGCGGAGG GACGATGATG AGCCTGGCCC TTGGACGAAG 1860  
 GAGGGGAAGG ACCGGATCAA GAAGCTGGGG CTGACCATGC AGTATCCAGA AGGCTACCTG 1920

WO 02/086443

GAAGCCCTGG CCAACCGAGA GCGAGAGAAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG 1980  
 CAGGAGGGGG GCTTCGCGTC CCCCAGGACG GCGAAGGGCA AGTGGAAAGC GAAGTCGGCA 2040  
 GGAGGTGGCC CGAGCAGGGC CGGGTCCCCG CGCCGGACAT CCAAGAAAAC CAAGGTGGAG 2100  
 CCTACAGTCT TCACGGCCCA GCAGAGCAGC CTCATCAGAG AGGACAAAGC CAACGCCAAG 2160  
 CTGTGGAATG AGGTCCCTGC GTCACTCAAG GACCAGCCGG CGAGCGGCAG CCCGTCCAG 2220  
 TTGTTCTCTGA GTAAAGTGGG GGAGACGTTT CAGTGTATCT GCTGTGAGGA GCTGGTGTTC 2280  
 CGGCCCATCA CGACCGTGTG CCAGCACAAAC GTGTGCAAGG ACTGCCTTGA CAGATCCTTT 2340  
 CGGGCACAGG TGTTCAGCTG CCTCGCTGCG CGCTACGACC TGGGCCGCGC CTATGCCATG 2400  
 CAGGTGAACC AGCCTCTGCA GACCGTCTCT AACCGCTCT TCCCGGCTA CGGCAATGGC 2460  
 CGGTGATCTC CAAGCACTTC TCGACAGGCG TTTTGTGTA AACGTGTGG AGGGCTCGTT 2520  
 CATCGGCACT GATTTTGTTC TTAGTGGGCT TAACTTAAAC AGGTAGTGT TCCTCCGTTT 2580  
 CCTAAAAAGG TTTGCTCTTC TTTTCTTTTA TTTTATTTT TCAATCTAT ACATTTTCAG 2640  
 GAATTTATGT ATTCTGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTTGT TCTTTGAAAA 2700  
 CATAAAGACC TGAATTTCT CGACAAAACA ACACAAGATT TTTTAAAGAT GGAATCAGAA 2760  
 ACTACGTGGT GTGGAGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAAG TTGCTGCCAC 2820  
 CAACTCTTTA AGAAGCGGAC AGGATCAGTC CTCTCTAGG GTTCTGGCCC CCAAGGTGAG 2880  
 AGCAAGCATC TTCTTGACAG CATTTTGTCA TCTAAAGTCC AGTGACATGG TTCCCGTGG 2940  
 TGGCCCGTGG CAGCCCGTGG CATGGCGTGG CTCAGCTGTC TGTGGAAGTT GTTGCAAGGA 3000  
 AAAGAGGAAA CACTCTGGGC CTAGTTCAAA CCTTTGCGCT AAAGCCATCC CCCACCAGAC 3060  
 TGCTTAGGCT CTGAGATCCG CGTGAAGAGT CCTCTGCCCA CGAGAGCAGG GAGTTGGGGC 3120  
 CAGCGAGAAA TGGCCTCAAG GGGACTCTGC TCCAGTGGG GCCAGGCGTG TGACTGACGC 3180  
 TGTCCGACGA AGGCGGCCAC GAGCGGACGC CAGCACAGA AGTCACTGC AAGTGCCCTT 3240  
 GATTCGTCTC TTCTTTCTAA AGACGACAGT CTTTGTGTT AGCACTGAAT TATTGAAAA 3300  
 GTCAACGAGA TTCTAGAAAC TGGGTCTATC CAGTTCTTCC TGACACCGGA TGGGTGCTTG 3360  
 GGAACCGTTT GAGCCTTATA GATCATTTAC ATTCAATTT TTTAACTCAG CAAGTGAGAA 3420  
 CTTACAGAGG GGTTTTTTTT TAATTTTTTT TTCTCTTAAT GAACACATTT TCTAAATGAA 3480  
 TTTTCTTGT AGTTACTGTA TATGTACCAA GAAAGATATA ACGTTAGGGT TTGGTGTGTT 3540  
 TTGTTTTTGT ATTTTCTTTC TTTTGAAGG GTTTGTAAAT TTTTCTAATT TTACCAAAGT 3600  
 TTGACGCTTA TACCTCAATA AACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACTG 3660  
 GAGCAATGTT ATTTTAAAG GGTTTTTTTC ACCTCCTTAT TCTTAGATTA TTAATGTATT 3720  
 AGGGAAGAA GAGACAAATT TGTGTAGGCT TTTTCTAAG TCCAGTACTT TGTCAGATT 3780  
 TTAGATTCTC AGAATAAATG TTTTTCACAG ATTGAAAAA AAAAAAAA

Seq ID NO: 455 Protein sequence  
Protein Accession #: NP\_037414.2

1 11 21 31 41 51  
 MWIQVRTMDG RQHTVDSLS RLTKVBELEK KIQELFHEVP GLQLRFYRKG QMEDGHTLED 60  
 YEVRLENDTIQ LLVRQSLVLP HSTKERDSEL SDTDSGCCLO QSESDKSSTH GEAAAEETDSR 120  
 PADEDMWDET ELGLYKVNVE VDARDTNMGA WFAQVVRVT RKAPSRDEPC SSTSRPALEE 180  
 DVIYHVYDD YPENGVVQMN SRDVRARART IIKWQDLEVG QVVMATYNPD NPKRGFPWD 240  
 AEISRKRETR TARLYANVV LGDDSLNDCR IIFVDEVFKI ERPGEGSPMV DNPMMRKSFP 300  
 SCKHKDDVN RLRCVCACHL CGGRQDFDKQ LMCDECDMAF HIYCLDPLPS SVPSSEDEWYC 360  
 PECRNDASEV VLAGERLRES KKKAKMASAT SSSQRDWGKG MACVGRTEK TIVPSNHYP 420  
 IPGIPVGTMM RPRVQVSESG VHRPHVAGIH GRSDNGAYSL VLAGGYEDDV DHCNFTYTG 480  
 GGGRLDGNK RTAQSCDQK LTNTNRALAL NCFAPINDQE GAERADWRSG KPVVRVRNVK 540  
 GGNKSKYAP EGNRYDGIYK VVKYWEKKG SGFLVWRYLL RRDDDEPGPW TREKDRICK 600  
 LGLTMQYPEG YLEALANRER EKENSKEKEE EQQEGGFASP RTGKGKWKRR SAGGGPSRAG 660  
 SPRTSKKTK VEPYSLIAQQ SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVEE 720  
 TFQICICQEL VFRPITTVQV HNVCKDCLDR SFRAQVFSFP ACRYDLGRSY AMQVNPQLQT

Seq ID NO: 456 DNA sequence  
Nucleic Acid Accession #: NM\_001200.1  
Coding sequence: 325..1514

1 11 21 31 41 51  
 GGGGACTTCT TGAATTGCA GGGGAAATAA CTTCGCGACC CCACTTTGCG CCGGTGCCCTT 60  
 TGCCCCACGG GAGCCTGCTT CGCATCTCC GAGCCCCACC GCCCCTCCAC TCCTCGGCCT 120  
 TGCCCGACAC TGAGACGCTG TTCCAGCGT GAAAGAGAG ACTGCGCGCG CGGCACCCGG 180  
 GAGAGGAGG AGGCAAGGAA AAGGAACGGA CATTGCGTCC TTGCGCCAGG TCCTTTGACC 240  
 AGAGTTTTTC CATGTGGAGC CTCTTTCAAT GGACGTGTCC CCGCGTCTCT CTTAGACGGA 300  
 CTGCGGCTCT CTAAGGTGCG ACCATGGTGG CCGGACCCCG CTGTCTCTTA CGGTGCTGTC 360  
 TTCCCGAGGT CCTCTGGGCG GCGCGGCTG GCCTCGTTCC GGAGCTGGGC CGCAGGAAGT 420  
 TCGCGCGCGC GTGTGCGGCG CGCCCTCAT CCCAGCCCTC TGACGAGGTC CTGAGCGAGT 480  
 TCGAGTTGCG GCTGCTCAGC ATGTTGCGCC TGAACAGAG ACCCAACCCC AGCAGGGACG 540  
 CGGTGCTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTGAG CCGGCTCAC 600  
 CGGCCCCAGA CCACCGGTG GAGAGGGCAG CCAGCCGAGC CAACACTGTG GCGAGCTTCC 660  
 ACCATGAAGA ATCTTTGGA GAACTACCAG AAACGAGTGG GAAAAACAAC CGGAGATTCT 720  
 TCTTTAATTT AAGTTCTATC CCCACGGAGG AGTTTATCAC CTCAGCAGAG CTTAGGTTT 780  
 TCGAGAACG GATGCAAGAT GCTTTAGGAA ACAATAGCAG TTTCCATCAC GGAATTAATA 840  
 TTTATGAAT CATAAAACCT GCAACAGCCA ACTCGAAT CCCCCTGACC AGACTTTTGG 900  
 ACACCAAGTT GGTGAATCAG AATGCAAGCA GGTGGGAAG TTTTGAATGTC ACCCCGCTG 960  
 TGATGCGGTG GACTGCACAG GGACACGCCA ACCATGGATT CGTGGTGGAA GTGCCCACT 1020  
 TGGAGGAGAA ACAAGGTGTC TCCAAGAGAC ATGTTAGGAT AAGCAGGTCT TTGACCAAG 1080  
 ATGACACAG CTGGTACAG ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAG 1140  
 GGCATCTCTC CCACAAAGA GAAAAACGTC AAGCCAAACA CAACAGCGG AAACGCTTCA 1200  
 AGTCCAGCTG TAAGAGACAC CCTTTGTACG TGGACTTCAG TGACGTGGGG TGGAAATGACT 1260  
 GGATTGTGGC TCCCGCGGGG TATCAGCGCT TTTACTGCCA CGGAGAAATG CCTTTCTCTC 1320  
 TGGCTGATCA TCTGAATCC ACTAATCATG CCATTGTGTA GACGTTGGTC AACTCTGTTA 1380  
 ACTTAAGAT TCCTAAGGCA TGCTGTGTC CCACAGAACT CAGTGTATC TCGATGCTGT 1440  
 ACCTTGACGA GAATGAAAG GTGTATTAA AGAATATCA GGACATGGTT GTGGAGGGTT 1500  
 GTGGGTGTCG CTAGTACAGC AAAATTAAT ACATAAATAT ATATATA

Seq ID NO: 457 Protein sequence  
Protein Accession #: NP\_001191.1

```

1      11      21      31      41      51
|      |      |      |      |      |
5  MVAGTRCLLA LLLPQVLLGG AAGLVPELGR RKPAAASSGR PSSQPSDEVL SEFELRLLSM 60
   FGLKQRPPTPS RDAVVPFMYL DLYRRHSGQP GSPAFDHRLE RAASRANTVR SFRHEESLEE 120
   LPETSGKTRR RPFPLNLSSIP TEEPITSABL QVFRQMQDA LGNNSPFHRH INIYEIKPA 180
   TANSKFPVTR LLDT

```

Seq ID NO: 458 DNA sequence  
 Nucleic Acid Accession #: NM\_001999.2  
 Coding sequence: 1..8736

```

15 1      11      21      31      41      51
   |      |      |      |      |      |
   ATGGGGAGAA GACGGAGGCT GTGTCTCCAG CTCTACTTCC TGTGGCTGGG CTGTGTGGTG 60
   CTCTGGGGCG AGGGCACGGC CGGCCAGCCT CAGCCTCTCT CGCCAAGCC GCCCGGGCCC 120
   CAGCCCGCGC CGCAACAGGT TGGTCCGCT ACAGCAGGCT CTGAAGCGCG GTTCTTAGCG 180
   CCGAGTATC GCGAGGAGGG TCCCGCAGTG CCGAGCCGCG TCCGCGGCG AGGACAGCAG 240
   GACGTCTCC GAGGGCCCAA CGTGTGGCG TCCAGATTCC ACTCCTACTG CTGCCCTGGA 300
20 TGAAGAGCGC TCCCTGGAGG AAACCACTGC ATTGTCCCGA TTGTAGAAA TAGTTGTGGA 360
   GATGGATTTT GTTCCCGTCC TAACATGTGT ACTTGTTCCT GTGGGCAAA ATCATCAACC 420
   TGTGGATCAA AATCAATTTA GCAATGTCAG GTGAGATGCA TGAATGGTGG GACCTGTGCA 480
   GATGACCACT GCCAGTGGCA GAAAGGATAT ATTGGAACCT ATTGTGGACA ACCTGTCTGT 540
   GAAAAATGAT GTCAAGATGG TGGACGTGTC ATGCCCAAAC CGTGTGCTTG TGTATTATGG 600
25 TTCACTGGTC CACAGTGTGA AAGAGATTAC AGGACAGGCC CGTGTTCAC TCAGGTCAAC 660
   AACCAAGATG GCCAAGGGCA GCTGACAGGC ATTGTCTGCA CGAAGACTCT GTGCTGTGCC 720
   AACCATGGAC GGGCGTGGGG CCATCCCTGT GAGATGTGTC CAGCCAGGCC TCAGCCCTGC 780
   CGACGGGGTT TCATCCCAAA CATCCGCACT GGAGCTTGCC AAGATGTGTA TGAATGCCAG 840
   GCTATCCAG GGAATGGCA AGGAGGAAAC TGTATCAATA CAGTGGGCTC TTTTGAATGC 900
30 AGATGCCCTG CTGGTCACAA ACAGAGTGAA ACTACTCAGA AATGTGAAGA CATTGATGAG 960
   TGCAGCATCA TTCTGGGATG ATGTGAAACT GGTGAATGTT CCAACACCGT GGAAGCTAT 1020
   TTTTGTGTTT GTCCAGCTGG ATATGTAAAC TCAACAGATG GCTCTCGATG CATCGATCAG 1080
   AGAACAGGCA TGTGTTCTC GGGCCTGGTG AATGGCCGCT GTGCACAAGA GCTCCGGGG 1140
   AGAATGACGA AATGCACTG CTGCTGTGAG CCTGGCCGCT GCTGGGGCAT CGGAACCAAT 1200
35 CTTGAAGCCT GTCTGTGAC AGGTTCAGAG GAATATCGCA GACTTTGATG GAATGACTT 1260
   CCAATGGGAG GAATTCAGG GAGTGTCTGT TCCAGACCTG GAGGCACTGG GGAATAATGG 1320
   TTTGGCCCAA GTGGCAATGG CAATGGCTAT GGGCCAGGAG GGACAGGCTT CATCCCATC 1380
   CCTGGAGGCA ATGGCTTTTC TCCTGGCGTT GGGGGAGCCG GTGTGGGGCG CGGGGACAG 1440
   GGACCTATCA TCACTGACT AACAAATCTG AACAGACAA TAGATATCTG TAAGCATCAT 1500
40 GCTAACCTTT GTTTAAATGG ACGCTGTATA CCAACTGTCT CAAGCTACCG ATGTGAATGC 1560
   AACATGGGTT ATAAGCAGGA TGCAAAATGA GATTGTATAG ATGTGTATGA ATGCACATCA 1620
   AATCCCTGCA CTAATGGAGA TTGTGTTAAC ACACCTGGTT CCTATTATTG TAAATGTCAT 1680
   GCTGGATTCC AGAGAGACTCC TACCAAGCAA GCATGCATTG ATATTGATGA GTGCATCCAG 1740
   AATGGGGTTC TTTGTAAAAA CGGTGATGTC GTGAACCTAG ATGGAAGTTT CCAATGCAAT 1800
45 TGCAATGCCG GCTTTGAATG AACTACAGAT GGAATAAAT GTGTGATCA TGATGAATGT 1860
   ACACTACCA ACATGTGTTT GAATGGAATG TGCATCAATG AAGATGGCAG CTTCAAGTGC 1920
   ATCTGCAAAC CAGGATTTGT CTGGCTCCA AATGGGCGTT ACTGTACTGA TGTGTATGAA 1980
   TGCCAGAGCC CAGGAATCTG CATGAATGGG CACTGCATCA ACAGTGAAGG GTCTTCGCG 2040
   TGTGACTGTC CCCCAGGCGT GGCTGTGGGC ATGGATGGAC GTGTGTGTGT TGATACTCAC 2100
50 ATGGCGAGTA CTGCTATGG AGGAATCAAG AAAGAGTGT GTGTGCGTCC TTTCCCGGCT 2160
   GCAGTGACCA AGTCCGAATG CTGCTGTGCC AATCCAGACT ATGTTTTTGG AGAACCTGTC 2220
   GACCCATGCC CTGCAAAAAA TTCAGCTGAA TTCCACGCCC TTTGTAGTAG TGGAGTAGT 2280
   ATCACTGTGG ATGGAAGAGA TATCAATGAA TGTGCTTTGG ATCCTGATAT ATGTGCCAAT 2340
   GGGATTGTTG AAAACTTACG TGTGTAGTAC CGTTGTAATT GCAACAGTGG CTATGAACCA 2400
55 GATGCTCTG GAAGAACTG TATTGACATT GATGAATGTT TAGTAAACAG ACTGCTTTGT 2460
   GATAACGGAT TGTGCCGAAA CAGCCAGGGA AGTTACAGCT GTACGTGCCC ACCAGGGTAT 2520
   GTGTTCAAGA CTGAGACAGA GACCTGTGAA GATATAAATG AATGTGAAG CAACCCATGT 2580
   GTCATGAGG CTTGAGAGAA CAACCTTGGA TCTTTCAATT GTGAATGTTT GCCCGGCAGC 2640
   AAACCTCAGT CCACAGGATT GATCTGTATT GACAGCCTGA AGGGGACCTG TTGGCTCAAC 2700
60 ATCCAGGACA CCGCTGTGTA GGTGAATATT AATGGAGCCA CTCTGAAATC TGAATGCTGT 2760
   GCCACCTCG GAGCGGCTG GGGGAGCCCC TGTGAGCGGT GTGAACCTAG TACAGCTTGC 2820
   CCAAGAGGSC TTGCCAGGAT TAAAGGTGTT ACGTGTGAAG ATGTTAATGA GTGTGAGGTG 2880
   TTCCCTGGCG TTTGTCCAAA TGGACGCTGT GTCAACAGTA AGGATCTTTT TCATTGGCAG 2940
   TGCCCTGAAG GCCTTACGTT GGATGGGACT GGGCGTGTAT GTTTGGAAT TCGCATGGAG 3000
65 CAGTGTACT TGAAGTGGGA TGAAGATGAA TGCATCCACC CCGTTCTCTG AAAGTTCCGC 3060
   ATGGAATGCT GCTGCTGTGC TGTGGGGCG GCTTGGGGCA CCGAGTGTGA GGAGTGCCCC 3120
   AAACCTGGCA CCAAGGAATA CAGACACTG TGCCCCCGCG GGGCTGGCTT TGCTAACCGA 3180
   GGGGATGTTT TTAAGTGGCG GCCATTTTAC AAAGACATCA ATGAATGCAA AGCATTTCTT 3240
   GGGATGTGCA CTTATGGGAA GTGCAGAAAT ACAATCGGAA GCTTCAAATG CCGTTGCAAT 3300
70 AGTGGCTTTG CTCTAGACAT GGAGGAAAGA AACTGCACGG ACATGACGTA GTGCAGGATT 3360
   TCTCCTGACC TCTGTGGCAG TGGAACTGTC GTCAATACAC CCGGCAGCTT TGAGTGCGAG 3420
   TGCTTCGAAG GCTATGAAAG TGGCTTCATG ATGATGAAGA ACTGCATGGA CATTGAAGGA 3480
   TGTGAACGTA ACCCTCTCCT TGTAGGGGT GGCACCTGTG TGAACACTGA GGGCAGCTTT 3540
   CAGTGTGACT GCCCACTGGG ACACGAGCTG TCACCATCCC GTGAGGACTG TGTGATATT 3600
75 AATGAATGCT CCGTGAGTGA CAATCTCTGC AGAAATGGAA AATGTGTGAA CATGATTGGA 3660
   ACCTATCAGT GCTCTTGCAA TCCTGGATAT CAGGCTACGC CAGACCGCCA GGGCTGTACA 3720
   GATATTGATG AATGTATGAT AATGAACGGA GGCTGTGACA CCGAGTGAC AAATTCAGAG 3780
   GGAAGCTACG AATGCAGCTG CAGTGAGGAT TATGCCCTGA TGCCAGATGG GAGATCGTGT 3840
   GCAGACATTG ATGAATGTGA AAACAATCCT GATATCTGTG ATGGGGGCA GTGTACCAAC 3900
80 ATTCTCTGGG AGTATCGCTG CCTCTGCTAT GATGGCTTCA TGGCTTCCAT GGACATGAAA 3960
   ACATGACATT ATGTCAATGA ATGTGACCTA AATTCAAATA TCTGCATGTT TGGGGAATGT 4020
   GAGAACACAA AGGATCCCTT CATTTGCCAC TGTGAGCTGG GTTACTCAGT GAAGAAGGGG 4080
   ACCACAGGAT GTACAGATGT GGATGAGTGT GAAATTTGGT CTCAATACCT GCACATGCAT 4140
85 GCCTCATGTC TGAATATCCC AGGAAGCTTC AAGTGTAGCT GCAGAGAAGG CTGATTGGA 4200
   AACGGCATCA AGTGTATTGA TCTGGACGAA TGTCTAATG GAACCCACCA GTGTAGCATC 4260
   AATGCTCAGT GTGTAATAC CCGGGCTCA TACCGCTGTG CTGTCTCGA AGGTTTCACT 4320
   GGTGATGCTT TTACTGTGCT AGATGTTGAT GAGTGTGAG AAAACATAAA CCTCTGTGAG 4380

```

	AACGGACAGT	GCCTTAATGT	CCCGGGTGCA	TATCGCTGCG	AGTGTGAGAT	GGGCTTCACT	4440
	CCAGCCTCAG	ACAGCAGATC	CTGCCAAGAT	ATTGATGAAT	GCTCCTTCCA	AAACATTGTG	4500
	GTCTCTGGAA	CATGTAATAA	CCTGCCTGGA	ATGTTTCATT	GCATCTGCGA	TGATGGTTAT	4560
5	GAATTGGACA	GAACAGGAGG	GAACGTGTACA	GATATTGATG	AGTGTGCAGA	TCCTATAAAC	4620
	TGTGTCAATG	GCCTATGTGT	CAACACGCCT	GGTGCCTATG	AGTGTAACTG	CCCCCCCGAT	4680
	TTTCAGTTGA	ACCCAACTGG	TGTGGGTGTG	GTGACCAACC	GTGTGGGCAA	CTGCTACCTG	4740
	AAGTTTGGAC	CTGGAGGAGA	TGGGAGTCTG	TCTTGCAACA	CCGAGATCGG	GGTGGGGCTC	4800
	AGTCGCTCTT	CATGCTGCTG	CTCTCTGGGA	AAGGCCTGGG	GAAACCCCTG	TGAGACATGC	4860
10	CCCCCTGTCA	ATAGCACTGA	ATATTACACC	CTGTGTCCCG	GAGGTGAAGG	CTTCAGACCT	4920
	AAACCCCATCA	CAATCATTTT	AGAAGACATT	GACGAATGCC	AGGAGTTACC	AGGTCTCTGC	4980
	CAGGGTGGAA	ACTGCATCAA	CACCTTTGGG	AGCTTCCAGT	GTGAGTGCCC	ACAAGGCTAC	5040
	TACCTCAGCG	AGGATAACCG	CATCTGTGAG	GATATTGATG	AGTGTTTTGC	ACATCCTGGT	5100
	GTGTGTGGGC	CTGGGACCTG	CTATAACACC	CTGGGAAATT	ACACCTGCAT	TTGCCCACTT	5160
15	GAGTACATGC	AGGTCAATGG	AGGCCACAAC	TGCATGGACA	TGAGAAAAGG	CTTTTGCTAC	5220
	CGAAGCTATA	ATGGAACCCAC	TTGTGAGAAT	GAGTTGCCCT	TCAATGTGAC	AAAAAGGATG	5280
	TGCTGCTGCA	CATATAATGT	GGGCAAGGCT	GGGAACAAC	CTTGTGAACC	ATGCCCAACT	5340
	CCAGGAACAG	CTGACTTTAA	AACCATATGT	GGAAATATTC	CTGGATTAC	TTTGACATT	5400
	CACACAGGAA	AAGCTGTGTA	CATTGATGAA	TGTAAAGAGA	TTCCAGGCAT	TTGTGCAAAAT	5460
20	GGTGTGTGCA	TTAACCAAGT	TGGCAGTTTC	CGCTGTGAAT	GCCCTACAGG	ATTCACTTAC	5520
	AATGACCTGC	TGTTGGTTTT	TGAAGATATA	GATGAGTGCA	GCAATGTGTA	TAATCTCTGC	5580
	CACGGGATGG	CAGACTGCAT	CAATAGTCCT	GGTAGTTACC	GCTGTGAATG	TGCCCGGGGT	5640
	TTCAAACATT	CACCCAAATGG	GGCCTGTGTA	GATCGCAATG	AATGTTTAGA	AATTCCTAAC	5700
	GTGTTGCAGT	ATGGCTTTGT	TGTTGATCTG	CAAGGAAGTT	ACCAGTGCAT	CTGCGACAAT	5760
25	GGCTTTAAGG	CTTCTCAGGA	CCAGACCATG	TGCTGGAATG	TTGATGAGTG	CGAGCGGCAC	5820
	CCATGTGGAA	ATGGAACCTG	TAAAAACACC	GTGGATCCTT	ATAACTGTCT	GTGCTACCCA	5880
	GGGTTTGAAC	TCACTCATAA	TAATGATTGC	CTGGACATAG	ATGAGTGCAG	TTCTTTTTTT	5940
	GGTCAGGTGT	GCAGAAATGG	ACGTTGTITT	AATGAAATTG	GTCTTTTCAA	GTGCTATGTC	6000
	AACGAAGGTT	ATGAACCTTAC	CCCAGATGSC	AAAACTGTA	TAGACACTAA	TGAGTGTGTC	6060
30	GGCCTTCCCG	GCTCTTGCTC	TCTTGTTACC	TGTCAGAAAT	TGGAGGGATC	CTTCAGATGC	6120
	ATCTGTCCCC	CAGGGTATGA	AGTAAAAAGC	GAGAACTGCA	TTGATATAAA	TGAATGTGAT	6180
	GAAGATCCCA	ACATTTGTCT	TTTTGGTTCC	TGTACTAATA	CTCCAGGGGG	CTTCAGTGC	6240
	CTCTGCCCCC	CTGGCTTTGT	ACTATCTGAT	AATGGACGGA	GATGCTTTGA	TACTCGCCAG	6300
	AGCTTCTGCT	TCACAAATTT	TGAAATGGA	AAGTGTCTG	TACCCAAAGC	TTTCAACACC	6360
35	ACAAAAGCAA	AATCTGCTG	TAGTAAGATG	ACAGGAGAGG	ACTGGGGGGA	CCCTGTGAG	6420
	CTGTGCCCCA	AAGACGATGA	AGTTGCATTT	CAGGATTTGT	GTCCATATGG	CCATGGAAC	6480
	GTCCCTAGTC	TTCAATGATC	ACGTGAAGAT	GTCAATGAGT	GTCTTGAGAG	CCAGGCATT	6540
	TGTTCAAATG	GTCAATGTAT	CAACACCGAC	GGATCTTTTC	GCTGTGAATG	TCCATGGGC	6600
	TACAACCTTG	ACTACACTGG	AGTACGCTGT	GTGGATCTG	ATGAGTGTTC	AATCGGCAAT	6660
40	CCGTGTGGAA	ATGGTACATG	CACCAATGTT	ATTGGGAGTT	TGGAATGCAA	TTGCAATGAA	6720
	GGCTTTGAGC	CAGGGCCCAT	GATGAATTTG	GAAGATATCA	ACGAATGTGC	CCAGAACCCA	6780
	CTGCTGTGCT	CTTTACGCTG	CATGAACACT	TTTGGGTCTT	ATGAATGCAC	GTGCCGATTT	6840
	GGCTATGGCC	TGAGGGAAGA	TCAAAAGATG	TGCAAAAGATC	TGGATGAATG	TGCTGAAGGG	6900
	TTACACGACT	GTGAATCTAG	GGGCATGATG	TGTAAAGATC	TAATCGGCAC	CTTCATGTGC	6960
45	ATCTGCCCTC	CTGGAATGGC	COGAAGGCC	GATGGAGAGG	GCTGTGTAGA	TGAAAATGAA	7020
	TGCAAGGACA	AGCCAGGAAT	CTGTGAAAT	GGACGTTGTC	TTAACATTAT	TGGAAGCTAT	7080
	AGATGTGAGT	GTAATGAAGG	ATTCCAGTCA	AGTTCTTCAG	GCACTGAATG	CCTTGACAAAT	7140
	CGACAGGGTC	TCTGCTTTGC	AGAGGTACTG	CAGACAATAT	GTCAAATGGC	ATCCAGTAGT	7200
	CGCAATCTCG	TCACCTAAGT	AGAAATGCTG	TGTGATGGTG	GGCGAGGCTG	GGGCACCAG	7260
50	TCCGAGCTTT	GCCCACTTCC	TGGAACCTGCC	CAGTACAAA	AGATATGTCC	TCATGGCCCA	7320
	GGATATACAA	CTGATGGAAG	AGATATTGAT	GAATGTAAGG	TAATGGCCAA	CCCTGTGCC	7380
	AATGGTCAGT	CTGCAATATC	CATGGGCTCA	TTCCGATGCT	TCTGCAAGGT	TGGCTACACC	7440
	ACAGACATCA	GTGGAACCTC	TTGTATAGAC	CTTGATGAAT	GCTCCCACTG	CCCGAAACCA	7500
	TGCAACTACA	TCTGCAAGAA	CACCTGAGGG	AGTTATCAGT	GTTCATGTCC	GAGGGGGTAT	7560
55	GTCCCTGCAAG	AGGATGGAAA	GACATGCAAA	GACCTTGATG	AATGTCAAAC	AAAGCAGCAT	7620
	AACTGCCAGT	TCCTCTGTGT	CAACACCCCTG	GGGGGGTTTA	CCTGTAAATG	TCCACCTGGT	7680
	TTTACACAGC	ATCACACTGC	TTGTATGAC	AACAACGAAT	GTGGGTCTCA	ACCTTTGCTT	7740
	TGTGGAGGAA	AGGGAATCTG	TCAAAACACT	CCAGGCAGTT	TCAGCTGTGA	ATGCCAAAGA	7800
	GGGTTCTCTC	TTGATGCCAC	CGGACTGAAC	TGTGAAGATG	TTGATGAATG	TGATGGGAAC	7860
60	CACAGGTGCC	AACAGCGCTG	CCAGAACATC	CTGGGTGGCT	ACAGATGTGG	CTGCCCCCAA	7920
	GGCTACATCC	AGCACTACCA	GTGGAATCAG	TGTGTGCGAT	AGAATGAATG	CTCCAATCCC	7980
	AATGCTGTGT	GCTCTGCTTC	CTGCTACAAC	ACCCTGGGGA	GTTCACAGTG	CGCCTGCCCC	8040
	TGCGGGTTCT	CCTTGCACCA	GTCTCCAGT	GCCTGCCACG	ACGTGAATGA	GTGCTGTGTC	8100
	TCCAAGAACC	CCTGCAATTA	CGGCTGCTCT	AACAAGGAGG	GGGGCTACCT	CTGTGGCTGC	8160
65	CCCCCTGGGT	ATTACAGAGT	GGGCAAGGC	CACCTGTGCT	CAGGAATGGG	ATTTAACAAG	8220
	GGGCAGTACC	TGTCAGTGA	TACAGAGGTC	GATGAGGAAA	ATGCTCTGTC	CCGAGAAGCA	8280
	TGCTACGAGT	GCAAAATCAA	CGGCTATCCT	AAGAAGAGCA	CGAGGCGAG	GAGAAGTATT	8340
	CATGAACCTG	ATCCCACTGC	TGTTGAACAG	ATCAGCCTAG	AGAGTGTGCA	CATGGACAGC	8400
	CCCGTCAACA	TGAAGTTCAA	CCTCTCCAC	CTCGGCTCTA	AGGAGCACAT	CCTGGAACCTA	8460
70	AGGCCCGCCA	TCCAGCCCTT	CAACAACAC	ATCCGTTATG	TCATCTCTCA	AGGGAACGAT	8520
	GACAGGCTCT	TCCGATCCA	CCAAAGGAAT	GGGCTCAGCT	ACTTGACAC	GGCCAGAAG	8580
	AAGCTCATGC	CCGGCACATA	CACACTGGAA	ATCACTAGCA	TCCCTCTCTA	CAAGAAGAAG	8640
	GAGCTTAAGA	AACTGGAAGA	GAGCAATGAG	GATGACTACC	TCCTAGGGGA	GCTTGGGGAG	8700
	GCTCTCAGAA	TGAGGCTGCA	GATTCAAGTC	TATTAACCGT	TCACAGACTT	GGGCCAGGC	8760
75	TCAAAATCCTA	GCACAGCCAG	TCTGCAGAAG	CATTGAAAA	GTCAAGGACT	AATTTTAAAG	8820
	AGGAAAAATA	ATAATAACTC	TTGTTTCTTT	CCTCCCTGTC	TTAGACTTTG	AATGTTGACC	8880
	CTCACAGGGA	GGGATAATTT	AGACTCTGGT	ATGGCCAAAG	ATTTAGAGCTC	AAAGGCAACC	8940
	TGCGTTACTG	TATTTTTTAT	ATAACTTCAT	TTTAAATAT	ATTTAAAGAA	ACCTAAATGT	9000
	TCAGATATAC	AGCATATGGC	ACTAAATGCA	CAAAAATAAT	GTGAGCTTTT	TTTTTTTTTT	9060
80	CCTGTTAGCA	GCTGTAAACA	CTTTGGGTAT	TTTGCTATAG	TTGCTAATTA	AAAAATATA	9120
	GATGTTTATT	TATTTTTAAT	GCAGTAATAT	ATGGAGAAAT	GAACAACTA	TGTAACAAA	9180
	AAGGGAAACT	CACCTGTTTT	TCCTTAGATT	TATAAATTG	AGCTATTTTT	TTTAGAGGTG	9240
	CTTTTTAAAA	ATCCAATAGA	TACAAGAGAT	GTTCCTTTG	GTTCCTGCCC	AGTCATCCAG	9300
	CTGATACACA	CCTGATCGAT	TTTAAAGAAA	GCCACACAGA	GCTGAATCGG	GCAGTGCTAA	9360
85	TCAATATTTT	AAAAGACATG	AATGTCATTA	GATCCTTTAT	AACGTAGATC	GAAGCCAAAG	9420
	CAGCTCATT	GTGACAACAT	TTCATATCAC	CAGACACACC	AGGCAACAGA	AGTTGAAGCA	9480
	CAACCACTGT	AGCAAAATAC	CTTGACTGCT	TGTGAGACCA	TTAGCATTGC	AGGCCAAACC	9540
	GTACTGTATT	TCCTTCTCAT	AACCTCAAGG	AACCATATGT	GCTACCCACA	ACACCTCATT	9600

WO 02/086443

PCT/US02/12476

CTTACCCAGG GTGCGCTGCG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC CGCCGTTCCC 9660  
TTGAAGGGA ACACCTGGCA TTCTGTGGTG TTTCTGCTGT TCTTAAATAA TGGTGCAATT 9720  
ATTATGTTCA AGTTATTCCA GGATTGCCAT ATGTGCAAA ACATCATGCA ATGCAGCCAA 9780  
GGAATATATG TTGTGTTTGT TGTITTTAAAC CCATTTTTTT TTTAGAAATT TCATTATATC 9840  
TGTAGTTATA CACCATATGC CTCATTTTAT CATAGCCTAT TGTGTATGAA AGATGTTTGT 9900  
ACAAATGAAT GATGTTTAGT TTGCTTTAGT CATTTAAAAA GATATTGTAC CAGGATGTGC 9960  
TATTAGAGAC ACGTATCCAT TATTTCTCTC AACCCAAAGAA CCTGTTTCTT GGACCAAGTA 10020  
CCAAACCTCA TATGTGAAT GGCCAAAGCA CATGCAGGCT CTGGTGTGTT CCTCTCAAC 10080  
CTGTGCTGAC CAAAGATTAG TAACCAAGTA TACCCAGTAT TTTGAGGTTT TATTGTTTTT 10140  
TTAATAACTA AAAAAAACT CGTGCC

Seq ID NO: 459 Protein sequence  
Protein Accession #: NP\_001990.1

1 11 21 31 41 51  
MGRRRRLCLQ LYFLWLGCVV LWAQGTAGQP QPPPPKPPRP QPPPPQVRSR TAGSEGGFLA 60  
PEYREGAAGV ASRVRRRGQQ DVLRGPNVCG SRPHSYCCPG WKTLPGGNQC IVPICRNSCG 120  
DGFCSPRNM TCSSGQISST CGSKSIQCCS VRCMNGGTCA DDHCQCKQGY IGTYCQGPVC 180  
20 ENGCGNGRC IAQPCACVVG FTGPQCERDY RTGFCPTQVN NQMCQQLTG IVCTRTLCCA 240  
TTGRANGHPC EMCPAQPPQC RRGFPINRT GACQDVDEQC AIPGICQGGN CINTVGSPEC 300  
RCFAGHKSSE TTQKCEDIDE CSIIIPGICET GECSNTVGSY FCVCPRGYVT STDGSRCIDQ 360  
RTGMCPSGLV NGRCAGQELPG RMTKMQQCCCE PGRCWGIGTI PEACPVRGSE EYRLCMDGL 420  
PMGIPGSAG SRPGGTGGNG PAPSGNGNGY GPGGTGFPII PGGNGFSPGV GGAGVGAGGQ 480  
25 GPIITGLTIL NQITIDICKH ANLCLNGRCI PTVSSYRCEC NMGYKQDANG DCIDVDECTS 540  
NPCTNGDCVN TPSSYYCKCH AGFQRTPTKQ ACIDIDECIQ NGVLCKNGRC VNSDGSFQCI 600  
CNAGFELTID GKNCDVHDEC TTNMCLANG CINEDGSFKC ICKPGFVLAP NGRYCTDVDE 660  
CQTGPICMG BCINSESGSP CDCPPGLAVG MDGRVCVDTH MRSTCYGGIK KGVCRPFPF 720  
30 AVTSECCCA NPDPYGPSPC QPCPAKNSAE FHGLCSSGVG ITVDGRDINE CALDPDICAN 780  
GICENLRGSY RCNCSGSGY DASGRNCIDI DECLVNRLLC DNGLCRNTPG SYSTCTCPGY 840  
VFRTESETCE DINECESNFC VNGACRNLG SFNCBCSPGS KLSSTGLICI DSLKGTCLNL 900  
IQDSRCBVNI NGATLKSBECC ATLGAANGSP CERCELDTAC PRGLARIKGV TCEDVNECEV 960  
PPGVCPNGRC VNSKGSFBC CPGSLTLDGT GRVCLDIRME QCYLKWEDEE CIHPVPKGR 1020  
MDACCAVCA ANGTCECECP KPGTKEYETL CPRGAGFANR GDLVLTGRPFY KDINECKAFP 1080  
35 GMCTYKRCN TIGSPKRCN SGFALDMEER NCTDIDECRI SPDLCSGIC VNTPGSFECE 1140  
CFEGYESGPM MKRNCMDIDG CERNPLLCRG GTCVNTGSGF QCDCLPHEL SPSRDCVDI 1200  
NECSLSDNLC RNSKCVNMIG TYQCSNPGY QATPDROGCT DIDECHIMNG GCDTQCTNSE 1260  
GSYECSCSGE YALMPDGRSC ADIDECBNP DICGGQCTN IPGEYRCLCY DGFMSMDMK 1320  
40 TCDIDNECDL NSNINMFGBC ENTGCSFICH CQLGYSVKGK TTGCTDVEDC SIGAHCMDH 1380  
ASCLNTPGSF KCSGREGWIG NGIKCIDLDE CSNGTHQCSI NAQCVNTPGS YRCACSEGFT 1440  
GDGFTCSVDV ECAENINLCE NGQCLNVPGA YRCECEMGFT PASDSRSCQD IDECSFQNIC 1500  
VSGTCNLPG MFHCICDDGY ELDRTGNGCT DIDECDPIN CVNGLCVNTP GRYECNCPD 1560  
FQLNPTGVGC VDNBVGNCYL KPGPRGDSL SCNTBIVGVV SRSSCCSLG KAWGNPCETC 1620  
45 PPNVSTBYTT LCPGEGSPFP NPITIILEDI DECQELPGLC QGGNCINTFG SPQCECPQGY 1680  
YLSDETRICE DIDECEAHFG VCGPGTCYNT LGNYTCICPP EYMQVNGGHN CDMRKSFCY 1740  
RSYNGTTCEN ELFPNVTKRM CCCTYNVGA GNPCEPCEPT PGTADEFTIC GNIPGFTFDI 1800  
HTGKAVDIDE CKRIPGICAN GVCINQIGSF RCECPTGFSY NDLLLVCEDI DECSNGDNL 1860  
QRNADCINSP GSYRCECAAG FKLSPNAGV DRNECLEIPN VCSHGLVDL QGSYQICHN 1920  
50 GFKASQDQTM CNDVDECEH PCNGTCKNT VGSYNCLCYP GFELTHNDC LDIDECSSFP 1980  
GQVCRNGRC ENCIDINECD EDNINCLFSG CTNTPGGFQC LCPGFLVSD NGRRCFDTRQ 2040  
ICPPGYEVKS ENCIDINECD EDNINCLFSG CTNTPGGFQC LCPGFLVSD NGRRCFDTRQ 2100  
SFCPTNFENG KCSVPKAPNT TRAKCCSKM PEGWGDPE LCPKDEVAF QDLCPYHGT 2160  
VPSLHDTRED VNECLSPGI CSNGQCINTD GSFRCECPMG YNLDTGVRC VDTDECSIGN 2220  
55 PCNGTCTNV IGSFECNNE GFEPGPMNC EDINECAQNP LICALRCMT FGSYECTCPI 2280  
GYALREDQIM LHDCEAG LHDCEAGMM CKNLIGTPMC ICPPMARRP DGEAGVDENE 2340  
CRTPGICEN GRVNIIGSY RCEBNEGFS SSSGTECLDN RQGLCAEVL QTIQCMASS 2400  
RNLVTKSECC CDGGRGWGHQ CELCPLGTA QYKKICPHGP GVTTDGRDID ECKVMPNLCT 2460  
NGQCINTMS FRCPCVKGYT TDISGTSCID LDECSQSPK CNYICKNTEG SYQCSCPRGY 2520  
60 VLQDEGKTK DLDECTQKH NQFLCVNTL GGFTCKCPFG FTQHTACID NNECGSQPLL 2580  
CGKGKICQNT GGFSCBQQR GFLDGLN CEDVDECDGN HRCQHGQNI LGGYRCGCPQ 2640  
GYIQHYQWQ CVDENECSNP NACGSASCYN TLAGSYKACP SGFSFDQPS ACHDVNECSS 2700  
SKNFCNYGCS AGACCTAAG PPGYRVGQG HCVSGMGFNK GQYLSLDETEV DEENALSPEA 2760  
CYECKINGYP KXDSRQKRSI HEPDPTAVEQ ISLESVDMDS FVMKFNLSH LGSKEHILEL 2820  
65 RPAIQPLNHH IRYVISQND DSVFRIHQRN GLSYLHTAKK KLMPTTYTLE ITSIPLYKKK 2880  
ELKKLEESNE DDVLLGELGE ALRMLQTL Y

Seq ID NO: 460 DNA sequence  
Nucleic Acid Accession #: NM\_013372.1  
Coding sequence: 63..617

1 11 21 31 41 51  
GCGGCCGAC TCAGGCCAC GCGTCGAAAG CGCAGGCCCC GAGGACCGCG CGCACTGACA 60  
GTATGAGCG CACAGCCTAC ACGGTGGGAG CCTGTCTCT CCTCTGSGG ACCCTGCTGC 120  
75 CGGCTGCTGA AGGGAAGAG AAGGGTCCC AAGGTGCCAT CCCCCGCCA GACAGGCCC 180  
AGCACAATGA CTCAGAGCAG ACTCAGTCG CCCAGCAGCC TGGCTCCAGG AACCGGGGC 240  
GGGCGCAAG GGGGGCACT GCCATGCCCG GGGAGGAGGT GCTGGAGTCC AGCCAGAGG 300  
CCCTGCTATG GACGAGCGC AAATACCTGA AGCAGAGCTG GTGCAAAACC CAGCGCTTA 360  
80 AGCAGACCAT CACAGAGAA GGCTGCAACA GTCGACCAT CATCAACCGC TTCTGTACG 420  
GCCAGTGCAA CTCTTCTAC ATCCCCAGGC ACATCCGAA GGAGGAAGGT TCCTTTCACT 480  
CCTGCTCTT CTGCAAGCCC AAGAAATCA CTACCATGAT GGTACACTC AACTGCCCTG 540  
AACTACAGCC ACTACCAAG AAGAAGAGAG TCACAGCTGT GAAGCAGTGT CGTTCATAT 600  
CCATCGATT GGAATAGGC AAATCCAGGT GCACCCAGCA TGTCTAGGA ATGCAGCCCC 660  
85 AGGAAGTCCC AGACCTAAA CAACAGATT CTTACTTGGC TTAACCTAG AGGCAGAGG 720  
AACCCTCAGC TGCCCTCTG CAGGAGCCTG CTTGTGCGTA GTTGTGTGTC ATGAGTGTG 780  
ATGGGTGCT GTGGGTGTTT TTAGACACCA GAGAAAACAC AGTCTCTGCT AGAGAGCACT 840  
CCCTATTGT TAAACATATC TGCTTTAATG GGGATGTACC AGAAACCCAC CTCACCCCG 900



CTCACATCTA AAGGGGCGGG GCGGTGGTCT GGTTCGACT TTGTGTTTTT GTGCCCTCCT 960  
 GGGGACCAAGA ATCTCCTTTC GGAATGAATG TTCAATGGAAG AGGCTCCTCT GAGGGCAAGA 1020  
 GACCTGTGTTT AGTGTCTGAT TCGACATGGA AAGTCTCTT TAACCTGTGC TTGCATCCTC 1080  
 CTTTCTCCTT CTTCTCTACA ATCCATCTCT TCTTAAGTTG ATAGTGACTA TGTCACTCTA 1140  
 5 ATCTCTTGTT TGCCAAGGTT CCTAAATTAA TTCACCTTAA CATGATGCAA ATGTTTTTCA 1200  
 TTTTGTGAAG ACCCTCCAGA CTCTGGGAGA GGCTGGTGTG GGCAAGGACA AGCAGGATAG 1260  
 TGGAGTGAGA AAGGAGAGGT GAGGGGTGAG GCCAATCAG GTCCAGCAAA AGTCAGTAGG 1320  
 GACATTGACG AAGCTTGAAA GGCCATACC AGAACACAGG CTGATGCTTC TGAGAAAGTC 1380  
 10 TTTTCTAGT ATTTAACAGA ACCCAAGTGA ACAGAGGAGA AATGAGATTG CCAGAAAGTG 1440  
 ATTAACCTTG GCCGTGCAA TCTGCTCAA CCTAACCA AACTGAAAC ATAAATACTG 1500  
 ACCACTCCTA TGTGCGACC CAAGCAAGTT AGCTAAACCA AACCAACTCC TCTGCTTGT 1560  
 CCTCAGGTG GAAAGAGAG GTAGTTTGA ACTCTCTGA TAGGGGTGGG AATTAATCAA 1620  
 AACCCKCGA GGCTGAAATT CCTAATACCT TTCTTTATC GTGTTTATG TCAGCTCATT 1680  
 TCCATCCAC TATTTCCCAT AATGCTCTG AGAGCCACTA ACTTGATTGA TAAAGATCCT 1740  
 15 GCCTCTGCTG AGTGTACCTG ACAGTAAGTC TAAAGATGAR AGAGTTTAGG GACTACTCTG 1800  
 TTTTAGCAAG ARATATTKTG GGGGTCTTTT TGTTTAACT ATTGTCAGGA GATTGGGCTA 1860  
 RAGAGAAGAC GAGGAGAGTA AGGAAATAAA GGGRATTGCC TCTGGCTAGA GAGTAAGTTA 1920  
 GGTGTAAATA CCTGTGAAA ATGTAAGGGA TATGACCTCC CTTCCTTTAT GTGCTCACTG 1980  
 20 AGGATCTGAG GGGACCTGAT TAGGAGAGCA TAGCATCATG ATGTATTAGC TGTTCATCTG 2040  
 CTACTGGTGG GATGACATA ACTATTGTAA CTATTGAGTA TTTACTGGTA GGCACGTGCC 2100  
 TCTGATTAAA CTGGCCCTAG TGGCAATGCG TACTTAGGAT TGATCTAAGG GCCAAAGTGC 2160  
 AGGGTGGGTG AACTTTATTTG TACTTTGGAT TTGGTTAACC TGTTCCTTTC AAGCCTGAGG 2220  
 TTTTATATAC AAACCTCCCTG AATACCTCTT TTGCTTGTG TCTTCTCAGC CTCTAGCCA 2280  
 25 AGTCCTAGT AATATGAAA ACAAACTG CAGACTGAG ATTCACTGCG CGATCAAGGC 2340  
 TCTGGCATTG AGAGAACCTC TGCAACTCGA GAAGCTGTTT TTATTGCTT TTTGTTTGA 2400  
 TCCAGTGCTC TCCATCTAA CAACATAACA GGAGCCATT CAAGGGCGGA GATATTTTAA 2460  
 ACACCCAAAA TGTGGGCTCT GATTTTCAA CTTTTAACT CACTACTGAT GATTCTCACG 2520  
 CTAGGCGAAT TGTGCAAA ACATAGTGTG TGTGTTTGT ATACACTGTA TGACCCCAAC 2580  
 30 CCAAACTCTT GTATTGTCCA CATCTCCA CAATAAGCA CAGAGTGGAT TTAATTAAGC 2640  
 ACACAAATG TAAGGAGAAA TTTTGGGGT GGGAGAGAG AAAAGGGAAA GAAGCTGAAA 2700  
 ATGTAAACC ACACAGGGA GGAATAATGA CATTGAGAC CAGCAACAC TGAATTTCTC 2760  
 TTGTTGTTT AACTCTGCA CAAGAATGCA ATTTGCTTAA TGGAGATGAC TTAAGTTGGC 2820  
 AGCAGTAATC TTCTTTTAGG AGCTGTGACC ACAGCTTGGC ACATAAGTGC AGATTTGGCT 2880  
 35 CAAGTAAGAG GAATTTCTC AACAATACT TCACTGGAT TACTCAGCAG GTAACCTACC 2940  
 TAAAGCATA TCACTAGCCA AAGAGGGAAA TATCTGTCT TCTTACTGTG CCTATATTAA 3000  
 GACTAGTACA AATGTGGTGT GTCCTCAAC TTTCAATGAA AATGCCATAT CTATACCATA 3060  
 TTTTATCGA GTCAGTGATG ATGTAATGAT ATATTTTTC ATTATTATAG TAGAATATT 3120  
 TTATGGCAAG ATATTTGTGG TCTTGATCAT ACCTATTAAA ATAAATGCCA ACACCAATA 3180  
 40 TGAATTTAT GATGTACACT TTGTGCTTGG CATTAAGA AAAAAACRA CATCTGGAA 3240  
 GTCGTGAAGT TGTTTTTGT TACTGTAGGT CTTCAGAGT AAGAGTGTAA GTGAAAAATC 3300  
 TGGAGGAGAG GATAATTTCC ACTGTGTGGA ATGTGAATAG TTAATGAAA AGTTATGGTT 3360  
 ATTTAATGTA ATTATTACTT CAAATCCTTT GGTCACTGTG ATTTCAAGCA TGTTCCTTT 3420  
 TTCTCCTTTA TATGACTTTC TCTGAGTTGG GCAAGAAGA AGCTGACACA CGTATGTTG 3480  
 45 TTAGAGTCTT TTATCTGGT AGGGGAAACA AATCTTGAC CCAGCTGAAC ATGTCTTCCT 3540  
 GAGTCAGTGC CTGAATCTTT ATTTTAAA TTGAATGTTT CTAAAGGTT AACATTCTA 3600  
 AAGCAATATT AAGAAGAGT TTAATGTTA TTTTGAAGA TTTACGATGC ATGTATACAA 3660  
 ACGAATAGCA GATAATGATG ACTAGTTCAC ACATAAGTC CTTTAAAGGA GAAATCTAA 3720  
 AATGAAAGT GGATAACAG AACATTATA AGTGATCAGT TAATGCCTAA GAGTGAAAGT 3780  
 50 AGTTCTATTG ACATCTCTCA AGATATTAA TATCAACTGC ATTATGATT ATGTCTGCTT 3840  
 AAATCATTTA AAAACGGCAA AGAATTATAT AGACTATGAG GTACCTGTCT GTGAGGAGG 3900  
 ATGAAGGGG AGTTGATAGT CTCATAAAC TAATTTGGCT TCAAGTTTCA TGAATCTGTA 3960  
 ACTAGAATT AATTTTACC CCAATATGT TCTATATAGC CTTTGTCTAA GAGCACTAA 4020  
 TAAATTAAC CTATCTTTC AAAAAAAA

Seq ID NO: 461 Protein sequence  
Protein Accession #: NP\_037504.1

1 11 21 31 41 51  
 60 MSRTAYTVGA LLLLLGLTLLP AAEKKKKGSQ GAIPPPDKAQ HNDSEQTQSP QQPSRNRGR 60  
 GQGRGTAMPG EEVLESSQEA LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRFGY 120  
 QCNSFYIPRH IRKEEGSFQS CSFCKPKKPT TMMVTILNCP E LQPPTKKKRV TRVKQCRCT 180  
 IDLD

Seq ID NO: 462 DNA sequence  
Nucleic Acid Accession #: Bos sequence  
Coding sequence: 1..2733

1 11 21 31 41 51  
 70 ATGAAAGTTG GAGTGTGTG GCTCATTCTT TTCTTCACCT TCACTGACGG CCACGGTGGC 60  
 TTCCTGGGGA AAAATGATGG CATCAAAACA AAAAAAGAAC TCATTGTGAA TAAGAAAAA 120  
 CATCTAGGCC CAGTCGAAGA ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAAG 180  
 75 GAGAAAGAG ATTTGAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCACAT 240  
 GGSCTAATTA GAATTATCAG AGCAAGGCT ACCACAGACT GCAACAGCCT GAATGGAGTC 300  
 CTGCAGTGA CCGTGAAGA CAGCTACACC TGGTTTCTC CCTCATGCTT TGATCCCCAG 360  
 AACTGCTACC TTCACAGGCG TGGAGCACTC CCAAGCTGTG AATGTCTAT CTCAACCTC 420  
 AGCCAGAGTG TCAATTTCTG TGAGAGAAAC AAGATTGGG GCACCTTTCA AATTAATGAA 480  
 80 AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAATA TGCAATGGA 540  
 ATTGAATTC AACCTAAAA AGCATATGAA AGAATTCAG GTTTTGAGTC GGTTCAGGTC 600  
 ACCCAATTTT GAAATGGAAG CATCGTTGCT GGGTATGAG TTGTTGGCTC CAGCAGTGCA 660  
 TCTGAATGCG TGTGAGCAT TGAACATGTT GCCGAGAAG CTAAGACAGC CCTTCACAG 720  
 CTGTTTCCAT TAGAAGACGG CTCTTTGAGA GTGTTTCGAA AAGCCAGTGT TAATGACATT 780  
 85 GTCTTTGGAT TTGGGTCCAA GGATGATGAA TATACCTGCG CCTGCAGCAG TGGCTACAGG 840  
 GGAAACATCA CAGCCAAAGT TGAGTCTCT GGGTGGCAG TCACTAGGGA GACTTGTGTG 900  
 CTCTCTCTGC TTGAAGAACT GAACAAGAA TTCAATATGA TTGTAGGCAA TGCCACTGAG 960  
 GCAGCTGTGT CATCTTCGT GCAAAATCTT TCTGTCATCA TTGCGCAAAA CCCATCAACC 1020



ACAGTGGGGA ATCTGGCTTC GGTGGTGTG ATTCTGAGCA ATATTTTCATC TCTGTCACTG 1080  
 GCCAGCCATT TCAGGGTGTG CAATTCAACA ATGGAGGATG TCATCASTAT AGCTGACAAAT 1140  
 ATCCTTAATT CAGCTCAGT AACCACTGG ACAGTCTTAC TCGGGGAAGA AAAGTATGCC 1200  
 AGCTCAGGT TACTAGAGAC ATTAGAAAAC ATCAGCACTG TGGTGCCTCC GACAGCTCTT 1260  
 CCTCTGAATT TTCTCGGAA ATTCATTGAC TGGAAAGGGA TTCCAGTGAA CAAAAGCCAA 1320  
 CTCAAAAGGG GTTACAGCTA TCAGATTAAA ATGTGTCCCC AAAATACATC TATTCCCATC 1380  
 AGAGGCCGTG TGTAAATTGG GTCAAGCCAA TTCCAGAGAT CCCTTCCAGA AACTATTATC 1440  
 AGCATGCCCT GGTGACTCT GGGGAACATT CTACCCGTTT CCAAAAATGG AAATGCTCAG 1500  
 GTCAATGGAC CTGTGATATC CACGGTTATT CAAACTATT CCATAAATGA AGTTTTCCTA 1560  
 TTTTTCCTCA AGATAGAGTC AAACCTGAGC CAGCCTCATT GTGTGTTTTG GGATTTTCAGT 1620  
 CATTTGCAGT GGAACGATG AGGCTGCCAC CTAGTGAAATG AAACCTCAAGA CATGTGACG 1680  
 TGCCAAATGA CTCATTGAC CTCTCTCTCC ATATTGATGT CACCTTTTGT CCCCTCTACA 1740  
 ATCTTCCCGT TGTGAAAATG GATCACTAT GTGGGACTGG GTATCTCCAT TGGAGTCTC 1800  
 ATTTTATGCC TGATCATCGA GGCCTTGTGT TGAAGCAGA TTAATAAAG CCAACCTCT 1860  
 CACACAGTC GTATTGTCAT GGTGAACATA GCCCTGTCCC TCTTGATTGC TGATGTCTGG 1920  
 TTTATTGTTG GTGCCAGAT GGACACACG GTGAACCTTT CTGGAGTCTG CACAGCTGCT 1980  
 GTGTCTTTTA CACACTTCT CTACCTCTCT TTGTCTTCT GGATGCTCAT GCTTGGCATC 2040  
 CTGCTGGCTT AGCGATCAT CCTCGTGTTC CATCACATGG CCCAGCATTT GATGATGGCT 2100  
 GTTGATTTT GCTTGGGTGA TGGGTGCCCT CTCATTATAT CTGTCAATAC CATGTCTGTC 2160  
 ACGCAACCTA GCAATACCTA CAAAAGGAAA GATGTGTGTT GGCCTTAATG TCCCAATGGA 2220  
 AGCAAAACCA TCCTGGCTTT TGTGTCCCT GCATGCTCTA TTGTGGCTGT GAACCTCTGT 2280  
 GTGTGCTGTC TAGTCTCAC AAAGCTCTGG AGGCCGACTG TTGGGGAAAG ACTGAGTCGG 2340  
 GATGACAAGG CACCAATCAT CCGCGTGGGG AAGAGCTCC TCATTCTGAC CCCTCTGCTA 2400  
 GGGCTCACTT GGGCTTTTGG AATAGGAACA ATAGTGAGCA GCCAGAACTC GGCTTGGCAT 2460  
 GTTATTTTGT CTTTACTCAA TGCATTCCAG GGATTTTITA TCTTATGCTT TGGAACTACT 2520  
 TTGGACAGTA AGCTGCGACA ACTTCTGTTT AACAAAGTGT CTGCCTTAAG TTCTTGGAG 2580  
 CAAACAGAAA AGCAAAACTC ATCAGATTTA TCTGCCAAAC CCAAAATCTC AAAGCCTTTC 2640  
 AACCCACTGC AAACAAGAGG CCATTATGCA TTTTCTCATA CTGGAGATTC CTCGACAAAC 2700  
 ATCATGCTAA CTCAGTTTGT CTCAATGAA TAA

Seq ID NO: 463 Protein sequence  
 Protein Accession #: Bos sequence

1 11 21 31 41 51  
 MKVGVWLNLIS FFTPTDGHGG FLGKNDGIKT KKELIVNKKK HLGPFVEEYQL LLQVTVRDSK 60  
 EKRDLRNLFLK LLKPPLLNWH GLIRILRAKA TDCNSLNGV LQCTCEDSYT WFPSCLDPO 120  
 NCVLHTAGAL PSCECHLNNH SQSVNFCERT KINGTFKINE RPTNDLLNSS SATYSKYANG 180  
 IEIQLKAYE RIGQFESVQV TQFRNGSIVA GYEVVSSSSA SELLSAIEHV AEKAKTALHK 240  
 LFPLEDGSFR VFGKAQNDI VFGFGSKDDE YTLPCSSGYR GNITAKCESS GWQVIRETCV 300  
 LSLLEELNKN FSMIVGNATE AAVSSPVQNL SVIIRQNPST TVGNLASVVS ILSNSSL 360  
 ASHFVRNSNT MEDVISIADN ILNSASVTNW TVLLREKYA SSRLLLETLEN ISTLVPTAL 420  
 PLNFSRKFID WKGIIPVNSQ LKRGYSYQIK MCPQNTSIFI RGRVLIGSDQ FQRLPETII 480  
 SMASLTLMNI LPVSKNGNAQ VNGFVISTVI QNYSINEVEL FFSKIESNLS QPHCFVWDFS 540  
 HLQWNDAGCH LVNETQDQIVT CQCHLTSTFS ILMSFPVPST IFPVVKWITY VGLGISIGSL 600  
 ILCLIIIEALF WKQIKSQSTS HTRRICMVNI ALSLLIADVW FIVGATVDIT VNPBGVCTAA 660  
 VFFTHFFYLS LFFWMLMGI LLAIRIILVF HEMAQHLMMA VGFCLGYGCP LIISVITIAV 720  
 TQPSNTYKPK DLWNLWNSG SKPLLAFFVP ALAIVAVNFV VVLLVLTCLN RPTVGERLSR 780  
 DDKATIIIRVG KSLLLILPLL GLTWGPGIQT IVDSONLAWH VIPALLNAFQ GPFILCFGIL 840  
 LDKLRLQLF NKLKLSLSSWK QTEKQNSDDL SAKPKFSKPF NPLQNKHYA FSHTEGSSDN 900  
 IMLTQPVSENE

Seq ID NO: 464 DNA sequence  
 Nucleic Acid Accession #: AB035089.1  
 Coding sequence: 9845..10219

1 11 21 31 41 51  
 GGGCATGCAG CCATCGGGGA AAATCCATAG TGCAGATAAA GCAAGGAGGA AGAAGAAGGA 60  
 CAGTTCATAG AAAAGGGAGA ACATCAATAT AGGATGTTTC TTAGCAATAG AAAAGAAGG 120  
 CCAAGAGGAA TTAGGAGAG AGTTATAAGA GATCAGCAAG GGGACAGGGT TAGATTGGT 180  
 TTGGTTTGA AGCATACAGT AAATATGATG TCTGTCCCTG GCAGTGTGG CAGAGTAGGA 240  
 AGGAGGAAGG GAGGCAAGAG ATAATATCAT TTTCTCTGTG CTCCAAGTGT ACTTACATAT 300  
 GAGACTATTT CCCTCTCTGC TTTTCAAACC TTACTGAGAT TGTTTTCCCT CATGAAAACC 360  
 AAGAAAGGAA AGCTAGTTAG TCTTGTCTG AGGTGTGTTCA ATGTATACAT ATCTATATCT 420  
 GTAGACAGAA TCCTTGGGAA TACAGTAATT GACATATATT CTGTATTGTG ATGCTTGA 480  
 AATCTCCTCC ACTAACCAGT TTCCCTATAG ATTGCCACAA GCACATAATA AGAAACAATA 540  
 AATAAAATGT TCTCTTGACT TTGTACTTAA ACAATGCTGA GAAAACCTTA CAGCCTTCAT 600  
 AAGGAAGTGA GGTCCAGGAA AATCTAGGAG ATATTCTTAA ACCAATCTAT AAGGCATTA 660  
 GTAATGACAG GATATTCTCT GAAAGTGTAA TTTCCCATG AGGATTGTIT TTTAATTTCT 720  
 GGATTCCTGG AGCCAATGAA GTTGGTGTAT GTTTATGAAA TATCAAGAGA CATAAGTTGG 780  
 CAAGTGTTCA TATGCAAAA CTTCTGGAA TTTCTGAGTT CTCTGTGGCA ATATATGACA 840  
 TCAGGATATG TCCAGTCTCA CACACCAGGA TATGTCTCTT CTAGCCTGTC TATCATGTC 900  
 TAGGAGAACT ATTTAGGAAC AGAAAAAAT GCCTGAAATG ATTTCTCATT TGAACCTATC 960  
 CAAGCTTTCT CTAATTTTAA GCAAACCTCT GGTCAATTTT AGTTAGTACC TTTCTTAA 1020  
 TTCACCTTTC AGGGCAAAAC TCCGTGCCCT AGACGTTTAT CCATAGTCTG AAATCTCTCT 1080  
 CCATAGATTG GTCCCCTGTA ACCCGGTTT GTCTCAGCTT GTTATCCTGT TTTTCTTCT 1140  
 CCTCCATTCC CAGGATGAGC TTGTTGCTTC TGTCTATGTA GACATTAGAT TCCTTTCTCT 1200  
 TGGTACCGGA GTAAATCCAT CCTACTCCAA TAGAGGAAGG TCCATTTTGT TCTATAGCG 1260  
 CTGGATGCGAG ACTCAGCTGA GAAGACCAAT ATTCAATTTT GGAATTTCTT ATCTCAGATA 1320  
 TTTCTCTTTC TTTCTTTTTC TTCTATCTTT GGATTTTATG TCCATCAAGC CCCCATTAGT 1380  
 CTATCTCCCG ACTTCAATCA GGGAACTTAT ACCTCTTAAA CTCAITCAGA GACTCAAAAC 1440  
 ATATATATTG ATACAGGAGA CCTAAGAAGA GCATGTCTTG GGGGTGAGG AAACAGGCAG 1500  
 GTGAGAAATT TCCAGATTGG AAACACAGCT TCCTTTCTCC CATCCAGCCC CTACTTTCAG 1560  
 CCTATGTGTT TCTGGCACCT TGTGTAGAT AAATCTCCCT TGACTTTGTG ATGTGCTGAG 1620  
 AAACAAACTC CAGGCTGCT GTTAAAGAG GCCCATGACA ATACCAAGTG TTGGGGAGAA 1680  
 TGTGGAGAAA TCAGAACTCT ATTCAAGGTC GGTGGAATG CACACTTTGT CAGAATTTCTA 1740

	TTGAGAAAGAG	TCTGGCAATT	CCTCAAAATG	TTAACTTGA	TTTACCATAT	GACCCAGCGA	1800
	TTTCATTAT	AGGTTTATAC	TCAAAAGAAA	TGAAGAAATA	TGCCATGCAA	AAAAATGTAC	1860
	ATGAAAGGTC	ACAACATCAT	TATTCATAAT	AGTAAAAGGA	TGGAACCAAC	ACAAATGTCC	1920
5	ATCAACTTAT	GATTAAAGAA	AATCTGGTCT	ATTCATAGAA	TGGAATATTA	TTGACCCACA	1980
	AAAAGGAATG	ATGTACTGAT	CCATGCAATG	ATGTGGACAA	ACCATGAAAA	TAACACTAGA	2040
	TTAAAGAAAG	CAGTCAAA	AGGACTTACT	GTATGATGCC	ATTTACCTGA	AATGTTTGGG	2100
	ATAGGCAAA	CCATAGAAAC	AGGAGGTAGA	TTCTGGTIT	CCAGGTCTC	CAGGAAGGGA	2160
	AGAATGAAGT	ACAAGATTTT	TTTTGGAGGT	AGTGAATGT	TTGTGGAATG	AGATCATGAT	2220
10	GATGATAGCA	CAACTTTGTG	AATATAATAA	AATCATTGAA	TTGTACAGTT	GAATTTATGG	2280
	TATATAAAT	ATATGTTAAT	AAAAAGGGGG	TCCACAAAAC	AAACAGCCCC	CCACTCTGGT	2340
	TGTCAGGAG	ATATTGGATT	AAATGGCCTT	GGACAAACAC	CCCTCTCCCT	GGCCACAGAC	2400
	ATCTTTCAGA	TTACAGATA	TTCCAGGGGA	AACACTGGAA	TGAGTCTGAA	GCCAGGTGCT	2460
	AAACAGAAAG	ACCATTGAGA	AATGTTGTGA	TCCTGACAGG	TCAAGCAATT	TATTTTTCGG	2520
15	CTTCATTTT	AAATGTAAAA	TTAGAAAGCT	GCCATTAAAA	ATGGCCCGTC	TGTTTCAATT	2580
	GCTCTTCTCA	GTGTGACGCT	GTTAACTCAA	TGTGTTAGTC	TGTTTTCATG	CTGCTGATAA	2640
	AAACATACCT	GAGACTGGCA	AGAAAAGAG	GTTTAATTGG	GCTTAGAGTT	CCACGTGATT	2700
	GGGAGGCGCT	CAGAAATACA	GTAGAGGGCA	AAAGTTATTC	TTACATGGTG	CTGCAAGAG	2760
	AAGATGAGGA	AGAAGCAAAA	GAAGAAACCC	CTGATAAAC	CATCGGATCT	CCTGAGGCTT	2820
20	ATTAACATAT	ATGAGAAATG	CACAAGAAAG	ACCGCCCCC	ATGATTCAAT	TACCTCTACC	2880
	TGGGTCCCTC	CAATAACATG	TGGAAATCT	GGTAGATACA	ATTCAAGTTG	AGATTGGGT	2940
	GGGAACACAG	CCAAACCAT	TCACCTCAGCA	AGGCAGATAA	CTTCTCTACT	GAGCCTATGC	3000
	AACAGAAAAC	CATCTGGGAT	GGTTGTAAGG	GGCACAGGAA	GTGACTGGTA	GGATCACTGC	3060
	CAAGCTGAG	CATCTCAGGAG	AAGGCAATAG	AATCCTATTC	TCCATAGTAT	GCTATAAGAT	3120
25	ACTGAAGTAC	ACTCTCTCAC	TATCTCTTTC	GACTTAGAAT	TAGCACTACA	TTCTTTGTTA	3180
	TACAGAAAA	TTACTAAGCA	AATTCATAGG	ATGACAAAA	CTTTCAGAAC	TGAAAAACAG	3240
	GAAATGTAG	CTTTTATGTT	CTTTGGTATT	CGAAGTATGC	CTAAAAGACA	ATGCAAAATC	3300
	CAAGAAAAGA	ATGGTGGGGT	TTTTGTTTGT	TTGGTTTGT	TTTTGTTTGA	CAGCTGGAGT	3360
	AGAATACAAA	GGGATGGAGT	TGAACCAAA	GAGAGGAAAT	TGGAATTTCTA	AACCTATTCT	3420
30	CATTGGCATT	AGAAGGCGAC	CTACATGTAT	TTACATGAG	CCGGTGACTG	CTGACTTGCA	3480
	TTCTTATTTT	TTCCCTATAG	ATTAAGAAAG	AGGTACAATG	GTAGAACTGT	AATCCTGTCC	3540
	TTTGTATATA	ATTTTCATAT	TCATAAAGGT	GAGTGTAGC	COGCTTGTA	AATCTGAAGT	3600
	TGAGTAACCT	CAATATCTAA	CCACAGAGGG	AAAGGCAGCA	AGAGGAGAGG	CATAAATTTA	3660
	GGATCTCACC	CTTCATTCCA	CAGACACACA	CAGCCTCTCT	GCCCACTCT	GCTTCTCTTA	3720
35	GGACACACAG	TAGAGGCTTC	AAGCCTCTCC	AGCTTAATAA	CATGAATTTAT	TTTTGAGAA	3780
	AATAATGATA	CTGTGTCTTA	TATCATGCAT	CTCCTGCATT	CTGTCTGATT	ATATTTTACT	3840
	TATTCTGCCA	GAGCAAAAT	AAAAATCCTA	TTTCATCTGA	TTTGTCTTTT	ATCTAAATTC	3900
	CTTAGTTCCA	AGTAAACCAA	GGCACTTTTA	GGAACACAGA	GGGAGAGTGC	CTTGACGCCA	3960
40	GAGAGTCTTG	AAGGAGATGT	CAGGGAACCA	TCTTAACAGC	TGGTTGGATG	TGATCCACAG	4020
	AGGTCTCTTG	TGACATTTCA	TTGTAAGGCC	ATCCTACCTA	GCTCTAGTGT	AACCAAGCAAT	4080
	GAAAGAAAGA	TAAAGAGGGT	CGATTACTTA	TTTACAATAG	TCTTTAAAA	CGTAGTTTTC	4140
	TAAGCCTTCT	AATTAGGACA	TTAATATATT	TAATATATGC	ACATTGTAGA	AAGATTGAAG	4200
	CGTTAAAAAT	AAGAGAAAAA	CTTTAAATGT	CAAAATCTCA	CAACCCAGAT	ATATCATTTT	4260
	TTTAAGAAAA	TTGTACTTACA	AAATACCAAT	CCATTATTTA	AAGTCATTCT	GACAGGAATC	4320
45	TGATGCTTTT	CCAGAGGTTT	CAGATCACAT	CGAGTTCAAC	ATGAATTCAC	TCAGTGAAGC	4380
	CAACACCAAG	TTCAATGTTG	ATCTGTTCCA	ACAGTTGAGA	AAATCAAAAG	AGAACAACAT	4440
	CTTCTATTTC	CCTATCAGCA	TCACATCAGC	ATTAGGGATG	GTCCCTCTAG	GAGCCAAAGA	4500
	CAACACTGCA	CAACAAATTA	GCAAGGTAGC	TATCAGCATC	ATTACGTTGT	CCTGTTGCAG	4560
50	TTTTTCTCTG	GTTCGCTCGG	CTAGCAGCCA	GATGGTAATA	GATGTGGTGG	TCTGATGGGT	4620
	AGCACAGGGG	GCTGTGACAG	AATTTCCATA	ACTGTGAGAC	CATGACCTTA	AACAGATCTT	4680
	TTGAGTAAAG	TTTTCTTGTC	COGCTTCATG	TCTCTCCAG	GTTCCTCACT	TTGATCAAGT	4740
	CACAGAGAAC	TGACAGAAAC	AAGCTGCAAC	ATATCATGTG	AGTCACAGAG	CACCTGATT	4800
	CAGCTTTAGA	TCCCTGAACA	GGTCATAGTT	TAAACCTGGA	ACTTCACAAA	AACTAAGAAA	4860
	AGGCCAGTTT	TAGGGAANA	CTTGGACACA	AAGATTGAGA	CATACAGAGT	GGGTTGGCAT	4920
55	TTCATGGCAC	ATAATTATTA	TTCTCATTTT	CTGCTTACT	AAAAGACAGT	CAGCACTGTA	4980
	CCTCAGAGCA	TAGGCTCTGA	TCAGGATAGG	CTGGGTTGAG	ACTCCAGCTT	TGCTCTTCAC	5040
	AAATGATGAA	TAGAGCAGG	ACACAACTGC	TCGGAGTCCC	AGTGAACCTCA	TCCAGAGAAA	5100
	CTAAGGGTAA	GAANAATCT	GACTCAATAC	ATGCAAAATC	ATGCAAAATG	TTACAACAGT	5160
	GCCTTGCCCA	TAAAGTTCAT	AATAAATGTT	ATTATTATTA	TAAAGTAGCT	ATAATTATAC	5220
60	TAATCATAAT	AATGTGAAAA	TAATTTAATT	TTCAATTGAGT	CATTAAATGAG	ATTCAAGGGA	5280
	ATAAGCACAA	GTCCAAAGTAT	ATTTTGGAAA	ATGATTGCTA	TGGAATATAT	TGGTTTAGAG	5340
	CCTTAATAGT	GCAAAATGCT	TTGCTGGAA	GTAGAAAGTT	CTAGATTTAA	ACAGGCTTAG	5400
	GTTCAAAAT	TGGCACTTCT	AATTTATGTC	TCTATAAACA	GGGTTTTTTT	CCCATTTCTC	5460
	TGAGCTTTCT	TGTGTTTCAT	TGAATTGAAC	TAAAGACTTA	GAGTTACCCA	TGTAAAGTCC	5520
65	TTAGCCATGG	ACCTGGCATA	CACCTCTCTT	ACGTGCAGAG	AATGAACCATC	ATGAGGAAG	5580
	AGCCACAGAT	CAGTCAATGT	GTCTCAACAG	ATAATAGCAC	CAACAGGTAT	AACAGGGCTT	5640
	CCTGGCATAA	TCTATTTTAA	ATATCCAACC	TTCAACATAC	TGCTATCTCT	GATGACTGTT	5700
	AGAAGTGAAG	TATGTCCTCT	GCCCAATAGG	AGCTGAGAGT	TTAAGTGGGA	AGCTAAACCT	5760
	AACCTTTTAA	ACCAACAAGG	AGAAAATCTA	CTGGTAGACA	CGGCTGCATC	TTAGTTTCAG	5820
70	AAGAGAAAAG	ATTGCAGTAC	GTTAGAGCAA	GAAGAAATTT	CTGGAAGAAG	TCAATATATA	5880
	GGTGGATTTT	GAAGGATATT	TGAGGTGAAA	TACACCAATT	ATCAGGGAAT	AACATCAAG	5940
	GTCTCAATG	AGACTACACG	CATTTAGGGA	CTGATCTAAC	AGACTTAGCA	TGGGTTTATG	6000
	ATTACATG	ATACAGCAAT	TGAATGATCT	CCTTTTGTGA	TGTTTGAAGG	TTGATAGGTC	6060
	AGGAAATGTT	CATCACCAGT	TTCAAAAGCT	TCTGACTGAA	TTCAACAAAT	CCACTGATGC	6120
75	ATATGAGCTG	AGATCGCCA	ACAAGCTCTT	OGGAGAAAAG	ACGTATCAAT	TTTTACAGGT	6180
	AAATTCACCT	GGCCTACCCA	CATTTTCAAT	GCACTCTGAT	GTCTGTGTCT	CTGAGTGGCC	6240
	AAATGGAAGA	AAGCAAGGCA	GATGAGCCTG	GCGACCCAG	GTGGAGAGCA	TTTACTCAGA	6300
	GTGCATTAGC	TCCATTTC	CAACTCTCCC	CCACTGGAGT	GTCCAGACCC	CCAACGATAC	6360
	ATCACTGAG	TGTGGATTGA	GGGATAATCT	TGTGATAAAA	GAGGAGGTTG	TGTAAATAG	6420
80	TGAGTAAAG	TAATAAGTAA	TAAGATACCA	TCGATAAACT	GGCACTGACT	CAGTCACTA	6480
	CGATACATCT	TGTGTTGAAA	TGTATGACTA	ATGGGATATT	ATTGGAATGG	GCAGGCTTGG	6540
	GTGAGTTCT	GAGAATAGTT	GAGGAAGTAC	CAGGAATAT	TGAATGCACA	GGATGAAAGA	6600
	CAAAAACAAA	GATCAGAAAC	ATCATGGTTA	AAATTAATGG	AGAGAAGTCT	GAGAAGCAAT	6660
	GAATCTCCTT	CAGGGAAGCC	TGCTCTGCAG	TTTGCARACC	ACAGCCTCTT	CTGCTTCTGC	6720
	CTTTTGCCAA	GATGATATTG	ACCTTCAGTG	ACCTCTTCTT	TGTGCCAGCC	CACATTCCCC	6780
85	TTTTCATTTG	CCTACATGAC	ACCTGTATAA	AAATATCCAT	GGACAGGAGA	TACTGCATCT	6840
	ATTCAGGGTC	TGGATTGAGC	TTACTGTTGT	TACAAATAG	TAGATTGGT	AATATATAGT	6900
	TACATAAAT	ACTCTTAAT	CCTACTTCTT	CCTTCATATC	TCAAGGAAT	ATTAGATGC	6960

	CATCAAGAAA	TTTTACCAGA	CCAGTGTGGA	ATCTACTGAT	TTTGCAAATG	CTCCAGAAGA	7020
	AAGTCGAAG	AAGATTAACT	CCTGGGTGGA	AAGTCAAACG	AATGGTAGGA	GAGCCACCCA	7080
	TTATAGAAAC	ACCTTTGAGA	AACCTATGCC	AGTGAGCCTT	GTGCTTGACA	CTGCATGGGG	7140
5	GAACAGGTGT	GGGGATTGAG	ATGGGTTTGC	AGGGAGGGCT	GAAGAGGGCA	CTCCAGATGA	7200
	AGGATTTGTC	CAAATGAATA	TGAAGAGAGC	CTAGGGGAGC	CAAGGAGGAA	ATCAGAGGAA	7260
	GCCAATTAGA	TGGAACACA	TCTGGAGAAT	TATTTGCTTA	TGGCCCTGCA	TGCAATAGC	7320
	TTTGTGGATC	CCCTGTCTCC	GCTCAGACCT	ATTTTGAGAT	CATATCCTTT	ACTTTAAATC	7380
	AGACTCAAA	TTTTATGATG	AATATTTAAT	AGAAAACATT	AGAAAGCGTC	TCTCGTCTCC	7440
10	TTTACTAATT	GGGAAACAAG	CAGCTCTCTG	GTAAATCACC	CTTTTGTCTC	TGAGCTGGAG	7500
	CTGCCTGGAT	CACATCTGTA	GCCAATGTGT	TCTGCAGGGA	TTATCACAGC	TCTCTTCCCC	7560
	ATCAAGGGCA	AAGAGCTTGA	CAAAGTCTCC	ATTCTACAGA	CATCTTTCTT	ACCTCCCAAC	7620
	TCTCATTTACA	GGCCAACTT	ACAGCAACTC	AACATGAGAG	TGAATAGGAA	GATACCCCCG	7680
	GAGTAGTGT	CTGACAGCAC	AGGACATGCG	TTTCATATTA	CAGAGCTCAA	GTCACTCATC	7740
15	CTAAATGCA	ATCAGGGCCT	CCTTCTCTG	AATGGGACC	CCGTAGTTAA	AAAAAATAA	7800
	AAGTAGAAG	AGGAGGGAGG	GAGAAAGGAA	AGACACATGT	TGGAAGAGTA	GACAAAATCA	7860
	GTTTATCACT	ATTCCAAATC	AGATGATTGG	AGACATTTCAT	ACACAGAGAA	CGTGAACCTC	7920
	TTCTCTATCA	CAAGAAAGTA	TGCTCTCCATC	AAGGGTAACT	TTATACGACT	GGAGCCTTGA	7980
	AGAAAGCTGC	ATCTGGTGAA	CCACTGGTCA	GTGAGTCTAA	CAATTCAAAG	ATCAAAGTCA	8040
20	GTGAGTCTCA	AGCAGGGATT	TGGGTCAATA	ATTAACGATC	AGTCACGAAC	ATTGTCAAAG	8100
	CATCTTCCAG	ACAAGCCATT	TGTAGCTTGT	GTAAAAGACT	CTTTTATTCT	TTCCCTTGCA	8160
	GAAAAAATTA	AAAAACCTAT	TCCTGATGGG	ACTAATTGGCA	ATGATACGAG	ACTGGTTCCT	8220
	GTGAAGCCAA	TCTATTTCAA	AGGGCAGTGG	GAGAAATAAT	TTAAAAAGAA	AAACACTAAA	8280
	GAGGAAAAAT	TTTGGCCAA	CAAGGTATTG	TCTATATTTT	ATTATATATG	TGTAATATGT	8340
25	TAATACATGG	AATGTTAAAC	ATTTCTGATG	GAATGTAAAC	TGATAAGTAA	AAAATAAAAA	8400
	TGTTTCATGT	CTGTTATTTT	GTGTTTATAC	TCITTATACT	TTATTTAGTT	AGGAATACCT	8460
	GAAAAACTAT	TGTTTCTAAC	TCATGGAAAT	CCTGGGTTAT	TTCTTAGAAG	AAGAAGGATG	8520
	TGTTGCTATC	TCATTAATAT	TATCTTTTTT	GTCTTGTGTT	TCACGTGTTA	TTTGTGGGAC	8580
	ACATTGATT	ATTGAGATT	ACATACAAAT	CTGTACAGAT	GATGAGGCCAA	TACAATTCTC	8640
30	TTAATTTTGC	CTGTCTGGAG	GATGTACAGG	CCAAGGTCTC	GGAAATACCA	TACAAAGGCA	8700
	AAGATCTAAG	CATGATTGTG	CTGCTGCCAA	ATGAAATCGA	TGGTCTGCAG	AAGTAAAGAA	8760
	CTGCACTCTA	CAACTCTTCC	TTCTACTGCC	GGACATTTTT	CCAAAGATAC	CAAGTTTAAA	8820
	CAAGGTAAAA	GCTTATGACC	GAGTTGCGCT	AAAATGATGA	AAAATTCTAA	ATGAGGAATG	8880
	ATGACTCACC	TTCATATTAT	AAATATTGGA	GCAATAGGCC	TGACACAAAC	TGAAAGCTTA	8940
35	GTTTTGTGTT	GTGTTTGTGT	TTTTATTATT	ATTATTATAA	TACTTTAAGC	TTTAGGGTAC	9000
	ATGTGCACAA	TGTGCAAGTT	AGTTACATAT	GTATACATGT	GCCATGCTGG	TGTGCTGCAC	9060
	CCATTAACTC	ATCATTAGAG	GTTAGGTATA	TCTCCTAATG	CTATCCCTCC	CCCCCTCCCC	9120
	CACCCCAACAA	CAGTCTCTAG	AGTGTGATGT	TACCTTCTCG	TGTCCAAAGT	TTCTCATTTG	9180
	TCAATTCCCA	TCTATATGTT	AATTCCATCT	ATGGCTTAGT	TAATGATTAA	TTTATTAGAG	9240
40	TTACATGCAT	TGGATATCAA	TTTGATGATA	TTATTATGCA	GCAATTAAAA	CTTGACTGGG	9300
	AGAAATATAT	ACCAATGTGA	GGAAAGTTTA	CAATAGGCC	GAGTAGAAAA	GGGAATACAA	9360
	ATTTAGGAAT	TTAGGGGAAT	ACAAATTAAT	AATTGCAATG	TGTACTAAAT	AATGTATACA	9420
	GAAAAATATG	ATGAGCCTAT	TAAAAATGGA	CACATGATGT	AGGCTGTTGG	CACAAGAAAT	9480
	AGTGATACAT	ACAGTTCATT	GTGTACAAAA	TAATGTAATC	ATATTTTACA	TGTGTATCAT	9540
45	ACAGTGTGAT	ACATACATAT	GTACACATAT	ACATATACGT	AAAAACATGA	TTCTGTTTTT	9600
	ACATACATGT	ATATACATAT	ACACATATAA	CCCAATGTAT	TTATATATTC	AGGACTCATA	9660
	TTTTACCTAT	TAGAATAATA	ATGTCTATTA	AAGTGAACCT	TCTGTATTTC	ACATTTATTG	9720
	CCAAATTAAC	GAATCTCCAC	ATAGTCAATT	CATTGTTAAG	GTGTATTAGA	GATCGACAGT	9780
	TAGTCATATC	AGTTTCTTTT	TTCCATTGTT	ATAGCTTGAA	GAGAACTCA	CTGCTGAGAA	9840
50	ATTGATGGAA	TGGACAAGTT	TGCAGAATAT	GAGAGAGACA	TGTGTGCAAT	TACACTTACC	9900
	TGGTTCAAA	ATGGAAGAGA	GCTATGACCT	CAAGGACACG	TTGAGAACCA	TGGGAATGGT	9960
	GAATATCTTC	AATGCGGATG	CAGACCTCTC	AGGCATGACC	TGGAGCCACG	GTCTCTCAGT	10020
	ATCTAAAGTC	CTACACAAGG	CCTTTGTGGA	GGTCACTGAG	GAGGGAGTGG	AAGCTGCAGC	10080
	TGCCACCGCT	GTAGTAGTAG	TCGAATTATC	ATCTCCTTCA	ACTAATGAAG	AGTTCTGTTG	10140
55	TAATCACCCCT	TTCTTATCTT	TCATAAGGCA	AAATAAGACC	AACAGCATCC	TCTTCTATGG	10200
	CAGATTCTCA	TTCCCATAGA	TGCAATTAGT	CTGTCACTCC	ATTAGAAAAA	TGTTCACTTA	10260
	GAGGTGTTCT	GGTAACTGTA	TGCTGGGCAA	CAACAGATTC	TCTTGGCTCA	TATTTCTTTT	10320
	CTATCTCATC	TTGATGATGA	TAGTCATCAT	CAAGAATTTA	ATGATTAAAA	TAGCATGCCT	10380
60	TTCTCTCTTT	CTCTTAATAA	GCCACATAT	AAATGTACTT	TTCCCTCCAG	AAAAATTTCC	10440
	CTTGAGGAAA	AATGTCCAAG	ATAAGATGAA	TCATTTAATA	CCGTGCTCTC	TAAATTTGAA	10500
	ATATAATTCT	GTCTCTGACC	TGTTTAAAT	GAACCAAAAC	AAATCATACT	TTCTCTTCAA	10560
	ATTTAGCAAC	CTAGAAACAC	ACATTTCTTT	GAATTTAGGT	GATACCTAAA	TCCTCTTAT	10620
	GTITCTAAAT	TTTGTGATTC	TATAAAACAC	ATCATCAATA	AAATAATGAC	ATAAATCAT	10680
	TTTTGCTTTA	CCTGTTTCTT	CTCTGGAAG	GGCAAGTGTC	CAGTTACACA	TAGGAAGAT	10740
65	AATTTAGAGA	TATATTAACT	ATATATAAAG	GAAAAATAAA	AACAGAGTAG	TTCATGATGA	10800
	GCCTGGAGTA	GAAGGCATAT	CCCAGAACAG	GAGGAGCCTT	GTAACCCACA	TAGGAACCTC	10860
	CTATTTTATG	CTAAAGGGAT	AAGAACTCA	TTACAGGCTT	TGATGGTTGT	TTGTCAAAGA	10920
	GGGCATATAA	ATTATCATAT	CCACATCTAG	AAAATACATC	TCTGGCTACG	CTGATATCAA	10980
	TGGATGCGAG	GAAAGAACAG	TGTGGTTACC	ATATATAAAT	TAGGAATACA	TTAGAGTATT	11040
70	GGGAGTGGAA	ATGGAGAGAA	AGAAAGAGCC	TGGGGGAATT	ATTTAGGAAA	TAATAGTTAC	11100
	AGAAAGACAT	CTAAGTTGCT	GACCTATCTG	ACTGGATGGA	TGGAAGAATA	TCTTGTCTTC	11160
	GAGAGAAAAA	AAGACTTTGG	GTTTAAATTT	GTACTTGATG	AATTAAGGTA	CTTTTAATAT	11220
	TCAAATGGAT	TTGCCCTGGCA	GGCACTTGAA	GATATTAGTC	TAAATCTCAG	AAACAGAATA	11280
	TGATCTGAAG	CTCTAAATTT	GTGATATTCA	ATATAAATAC	TTTAGAGTCA	TTGGGATAAA	11340
75	TATGGTAGTT	GTAGCTAAAA	GCAAAAATAA	GATACTAGGG	AGAAAGGATA	AAGTTAGAA	11400
	AAAGAAAGAT	CTAGAATTGA	CCTTGAAGTA	TATCAGCATG	TGTAAGATC	AGGAATTGAT	11460
	CATTTTATT	TTCCAGAAAG	TAGCTTTTCT	TAGGGTTCCA	TATTTACTCC	CATAGATTCT	11520
	TCCC						

Seq ID NO: 465 Protein sequence  
Protein Accession #: BAB21525.1

	1	11	21	31	41	51	
85	MNSLSEANTK	FMPDLFQQPR	KSKENNIFYS	PISITSLQMG	VLLGAKDNTA	QKISKVLHFD	60
	QVTENTTEKA	ATYHVRDSGN	VHQFQKLLT	EPNKSTDAYE	LKIANKLFG	KTYQFLQEYL	120
	DAIKFYQTS	VESTDFANAP	EESRRKINSW	VESQTNKIK	NLPFDGTIGN	DTTLVLVNAI	180
	YFKQWENK	KKENTKEEKF	WPKNTYKSV	QMRQYNSFN	FALLEDVQAK	VLEIPYKGD	240

Seq ID NO: 466 DNA sequence  
Nucleic Acid Accession #: NM\_001910.1  
Coding sequence: 50..1240

1	11	21	31	41	51	
GGAGAGAAGA	AAGGAGGGGG	CAAGGGAGAA	GCTGTGTTCT	GGACTCACAA	TGAAAACGCT	60
CCTTCTTTTG	CTGCTGTTGG	CTCTGGAGCT	GGAGGAGGCC	CAAGGATCCC	TTACCAAGGCT	120
GCCCCCTAGG	AGGCATTCGT	CCCTCGAAGA	GAGGTCTGGG	GCACGAGGCG	AGCTCTCTGA	180
GTTCCTGGAA	TCCCATAAAT	TGGACATGAT	CCAGTTTCAC	GAGTCTCTGT	CAATGGACAA	240
GAGTGGCCAA	GAACCCCTCA	TCAACTACTT	GGATATGGAA	TACTTCGGCA	CTATCTCCAT	300
TGGCTCCCCA	CCACAGAACT	TCACTTGTCAT	CTCTGCACAT	GGCCTCTCCA	ACCTCTGGGT	360
CCCCCTCTGT	TACTGCTCAG	CGGCAGCGCT	CACAGAGCAC	AGCAGATCTC	AGCCTTCCCCA	420
GTCACAGCA	TACAGCCACT	CAGGTCAATC	TGTTCCATAT	CAGTATGGAA	CCGGAGCTT	480
GTCGGGATC	ATTGGAGCCG	ACCAAGTCTC	TGTGGAAGGA	CTAACCCTGG	TTGGCCAGCA	540
TTTGTGAGAA	ATGTGTACAG	AGCAGGCCCA	GACCTTTTGT	GATGCAAGAT	TGTATGGAAT	600
GCTGGGCTGT	GGATACCCCT	CCTTGGCTGT	GGGAGAGAGT	ATCTCAGTAT	TTGCACACAT	660
GATGGCTAG	AGGCTGGTGG	ACTTGGCGAT	GTTTCTTGTC	TACATGACGA	GTAACCCAGA	720
AGGTGGTCCG	GAAGGCGAGC	TGATTTTGGG	AGCTACGAC	CATCTCCATT	TCTCTGGGAG	780
CTCGAATGG	GTCCCAGTCA	CCAAGCAAGC	TTACTGGCAG	ATTGCCTGG	ATAACATCCA	840
GGTGGGAGCG	ACTGTATATG	TCGTGCTCGA	GGGTCTGCCAG	GGCATTTGGG	ACACAGGGAC	900
TTCCCTCAT	ACTGCGCCCT	CCGACACGAT	TAAAGCAGCT	CAAAAGCCCA	TTGGGGCAGC	960
CCCCGTGGAT	GGAGAAATAT	CTGTGGAGTG	TGCCAACCTT	AAGCTCATGC	CGGATGTGAC	1020
CTTCACCAAT	AACGAGTGCT	CCATATCCCT	CAGGCCACAT	GGCTACACCC	TACTGACGAT	1080
CTTGGATGGA	ATGCAGTTCT	GCAGCAGTGG	CTTTCAGGAA	CTTGACATCC	ACCTCTCCAG	1140
TGGGCGCCCT	TGGATCTCTG	GGGATGTCTT	ACTCCGACAG	TTTATCTCAG	TATGTGACCG	1200
TGGGAATAAC	CTGTGGGAG	TGGCCCCAGC	AGTCCCTAAG	GGAGGGGGCT	TGTGTCGTGT	1260
CCTGCTGTCT	TGACACAGCT	TGAATATGTT	AGGCTGGGAC	ATTCTTTCAC	ACCAAAACAA	1320
GTTATTTTCC	AGACAATGTA	GCTGTTTCCA	GGGTGTCCGC	TGGAATTAAG	OCCTACAGAA	1380
GTATAGAAAT	ACAGACACAC	ACACACATAT	ACACACACAC	ACACTTCACA	CATACACACC	1440
ACTCCACACA	COGTCAATAT	GGAGGAAITTA	OGTTATACAT	TCATATTTTG	TATGTATTTT	1500
TGATTTATGAA	AATCAAAATAT	TTTCACATTT	GAITTATGAA	ATCTTCAAAC	ATATGCACAA	1560
GCAGAGATCA	TGTTATTAATA	AATCCCTTGG	CAACTCTTGC	CAGCCCTGAC	AACCCCTACA	1620
CACACGGCCA	GGCCTGTTTA	TCTACATCTG	TGCCCATCTC	TCTCTCCAGC	TCCCATGTCT	1680
GTACCTGGAT	CATTCTGAAG	CAAATTCGGA	GCATTACATC	ATTTTGTCCA	TAAATATTTT	1740
TAACTACCTT	AAATATACAA	TCGGAATTTCA	AGCATCTCCC	ATGTTCCCCC	AAATGTTTGG	1800
CTGTGTTTTT	AGTGTGATTT	TTTGTATTAG	GATTCAGACA	AGGCCCATAT	ATTGCATTTA	1860
TTTGAATGCT	CTGTAGTCTT	CTTTCATCTT	ACACAGATTT	GCATCATTTA	ACGTGTGCTG	1920
TGAAATATCC	GAGGTGTCACT	TGACATGTGT	TCTCTGAAAT	TACTTTTCTG	ATAAATGTGT	1980
AGTTAGATCT	GGAGGCTCTGA	TTTTGTGGCA	AAAATACTTC	CTAGGTGGTG	CTGGTACTTT	2040
CTGTGTGCAT	CCGTGTCAGA	GCGCAGATAAT	GCGTGGCCCT	CTCATAGTGT	AATGTTAAGA	2100
CTGCTGGGTT	GGTTTGGAGT	TCTTGGCTTT	AATCATCTTC	TACAAATGCT	AGCATTTT	

Seq ID NO: 457 Protein sequence  
Protein Accession #: NP\_001901.1

1	11	21	31	41	51	
MKTILLLLLLLV	LLELGEAQGS	LHRVPLRRHP	SLKKKLRARS	QLSEFWKSHN	LDMIQPTESC	60
SMDQSAREPL	INYLDMEXPT	TISIGSPGP	PQVIFDTGSS	NLWQVSVYCT	SPACKTHSRF	120
SPQSQSTYSQ	GGQSFSIQVG	TGSLSIGIIG	DVSEVGLTV	VGQVQFBSVT	EPGQTVDABE	180
FDGILGLGVY	SLAVGGQTVL	FDNMMAQNLV	DLPMFVSQYM	SNPEGGAGSE	LIFGGYDHS	240
FSGSLNWWVP	TRQAYWQIAL	DNIOQVGTVM	PCSEGCQAIM	DTGSLTITGP	SDKKIQLQNA	300
HGPAVDVDEG	AVECANLNVM	VDFTPTINGV	PYTLSTPAYT	LLDPVDGMQF	CSSGFGQLDI	360
IRPPAGPLWL	GDVPIROFVS	PVDGRGNVRG	LAPAVP			

Seq ID NO: 468 DNA sequence  
Nucleic Acid Accession #: NM\_018058.1  
Coding sequence: 319..1575

1	11	21	31	41	51	
TACGCGCTGC	GGGACCGGCA	GGGGAAAGCC	ATCGGGGTCA	CAGACTGCGA	CATCGACGGG	60
GACGGCCGGG	AGGAGATCTA	CTTCTCAAC	ACCAATTAAG	CCTTCTCGGG	GGTGGGACAT	120
TACACACGCA	AGTTGTTCAT	GTTCGCGAAT	AACCGGTATG	AAGACTACTCT	GAGGCATGAG	180
CTCAACGTGG	CCCGTGTGTT	GTTCCAGCTT	TTTGCGGGAC	GCTCTGTGGC	CTGTGTGGAC	240
AGAAAGGGCT	CTGACGACTA	CTCTATCTAT	ATTGCAATTT	ACGCTCATCGG	TAATGTGGGC	300
CTCGATGCCG	TTATGAATG	GGACCCCTGAG	CGCTGTGACC	TTCCCGGGGG	CAATCTGGGC	360
CTCAGAGATG	TGGCTGCTGA	GGCTGGGGTG	AGCAAATATA	CAGGGGGCCG	AGGGCTCAGC	420
GTGGGCGCCC	TCTCTCAGAC	CAGTGGCTCT	GATATCTTCT	GCGACAATGA	GAATGGGCCT	480
AACTTCTCTT	TCCTCAACCG	GGGGCATGGC	ACCTTTGTGG	ACGCTGCGGC	CAGTGTGTGT	540
GTGACGACAC	CCCAACGACA	TGGGCGGAGT	GTCCGCTCTG	CTGACTTCAA	CGCGATGAGC	600
AAAGTGAGCA	TGCTGTATCA	CAACTGGAAT	GGCCCCACG	GCGCTTATCT	GCAATGTAGC	660
ACCCATGGGA	AGGTCCGCTT	CCGGGACATG	GCCTTACCCA	AGTTCTCCAT	GCCTCTCCCT	720
GTCGCGCAGG	TCATCAGCCG	CGACTTTGAC	AATGACACAG	AGCTGGAGAT	CTTCTTCAAC	780
AACATTTGCT	ACCGCAGCTG	CTCAGCCAAC	GCCTCTTTC	GGCTCATCCG	TAGAGAGCAC	840
GGAGACCCCC	TGCTCAGAGA	GCTCAATCCC	GGCGACGCTC	TGGAGCTCTG	GGGCGGGGG	900
ACAGGGGGTG	TGGTGAACGA	TTTCGACGGA	ACAGGGATCG	TGGACTCTAT	CTTGTCCCAT	960
GGAGAGTCCA	TGGCTCAGCC	GCTGTCCGTC	TTTCGGGGCA	ATCAGGGCTT	CAACCAACAT	1020
TGGCTGCGAG	TGGTGGCCAC	CACCCGGGTT	GGGCGCTTTG	CCAGGGGAGC	TAGAGCTGTG	1080
CTCTACACCA	AGAGAGATGG	GGCCCACTG	AGGATCATCG	CCAGGGGGCT	AGGCTACCTG	1140
TGTGAGATGT	AGCCCGTGGC	ACACTTTGGC	CTCGGGAGAG	ATGAGACCCG	CATGTGTGGG	1200
GTGACGTGGC	GACATGTGCA	GATGTGTGAC	CGGAACGTGG	CACGGCGGGG	GAGTGAATCA	1260

GTGCTGAGAG TCCTCTACCC CGGGGATGAG GACACACTTC AGGACCCAGC CCCACTGGAG 1320  
 ACACCAATGA ATGCATCCAG TTCCCATTCG TGTGCCCTCG AGACAAGCCC GTATGTGTCA 1380  
 ACACCTATGG AAGCTACAGG TGCCGGACCA ACAAGAAGTG CAGTCGGGCG TACGAGCCCA 1440  
 ACCAGGATGG CACAGCTTCG GTGGGACTC TCGGCCAGTC ACCGGGCCCC CGCCCCACCA 1500  
 CCCCCACCGC TGTGCTGCC ACTGCGCGTG CTGCTGCCGC TGCTGGAGCT GCCACTGTCTG 1560  
 CACCGTCTCT CGTAGATGGA GATCTCAATC TGGGGTCGGT GGTAAAGGAG AGCTGCGAGC 1620  
 CCAGTGTCTG AGCAGGGGTG GGACATGAAC CAGCGGATGG AGTCCAGCAG GGGAGTGGGA 1680  
 AAGTGGGCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCTTGGGAG CTAGACCCCTC 1740  
 CCCAAGCCCA TCCATGCACA TTACTTAGCT AACAAATTAGG GAGACTCGTA AGGCCAGGCC 1800  
 CTGTGCTGGG CACATGACCTG TGATCAGAGC AGACAGGGTG GCTGCCCTGA TGGCGCTTAC 1860  
 ATTCAGTGGG GTCTAATGAC CATATCTTAG GACACAGATG TGCCAGGGA GGTGGTGTCA 1920  
 CTGCACAGGA AGTATGAGGA CTTTAGTGTG CTGAGTTCAA ATCTGATTG AGGAACCTCAC 1980  
 AAAGCTATGT GACCTTACAC CAGTCACTTA ACTTGTAGC CATCCATTAT CGCATCTGCA 2040  
 AAATGGGGAT TAAGAATAGA ATCTGGGGT TAGTGTGGAG ATTAGATTAA ATGTATGTAA 2100  
 GACACTTGGC ACAAACCTG GCACATAGTA AAGGCTCAAT AAAAAAAGT GCCTCTCACT 2160  
 GGGCTTTGTC AACACGTG

Seq ID NO: 469 Protein sequence  
 Protein Accession #: NP\_060528.1

1 11 21 31 41 51  
 MDPEASDLR GILALRDVAA EAGVSKYTGG RGVSVGPILS SSASDIPCDN ENGNPFLPHN 60  
 RGDGTPVDAA ASAGVDDPRQ HGRGVALADF NRDGKVDIVY GNWNGPHRLY LQMSTHGKVR 120  
 FRDIAGPKFS MPSPVRTVIT ADFDNDQELE IPFNNIAYRS SSANRLPRVI RREHGDPLIE 180  
 ELNPGDALEP EGRGTGGVVT DFDGDGMLDL ILSHGESMAQ PLSVFRGNQG FNNNWLVRVP 240  
 RTRVGAFARG AKVVLVTKKS GAHLRIIDGG SGYLCMEFV AHFGLGKDEA SSVVETWPDG 300  
 KQMSRVNASG EMNSVLEILY PRDETLQDP APLETFMNAS SSHSCALETS PVSFTPEAT 360  
 GAGPTRSAVG ATSPTRMAQP ANGLSASHRA PAPPPPPILL PLPLLLPLLE LPILLHRS

Seq ID NO: 470 DNA sequence  
 Nucleic Acid Accession #: AJ279016  
 Coding sequence: 1..1962

1 11 21 31 41 51  
 ATGTCCAGGA TGTACCGTT CCTGCTGCTG CTCTGGTTTC TGCCCATCAC TGAGGGGTCC 60  
 CAGCGGGCTG AACCCATGTT CACTGCAGTC ACCAACTCAG TTCTGCTCTC TGACTATGAC 120  
 AGTAATCCCA CCCAGCTCAA CTATGGTGTG GCAGTTACTG ATGTGAGCCA TGATGGGGAC 180  
 TTTGAGATCG TCGTGGCGGG GTACAATGGA CCCAACCTGG TTCTGAAGTA TGACCGGGCC 240  
 CAGAAGCGGC TGTGGAACAT CGCGGTGATG GAGCGCAGCT CACCCCTACTA CGCGCTGCGG 300  
 GACCGCGCAG GGAAGCCCAT CGGGGTCA CA GCTGCGACA TGACAGGGGA CGGCCGGGAG 360  
 GAGATCTACT TCCTCAACAC CAATAATGCC TTCTCGGGGG TGGCCACGTA CACCGACAAG 420  
 TTGTTCAAGT TCGCAATAA CCGGTGGGAA GACATCTGTA GCGATGAGGT CAACGTGGCC 480  
 CGTGTGTGGG CCAGCCTCTT TGCCGGAGCG TCTGTGGCCT GTGTGGACAG AAGGGCTCT 540  
 GGACGCTACT CTATCTACAT TGCCAATTAC GCCTACGGTA ATGTGGGGCC TGATGCCCTC 600  
 ATTGAAATGG ACCCTGAGGC CAGTGACCTC TCOCGGGCGA TTCTGGCGCT CAGAGATGTG 660  
 GCTGCTGAGG CTGGGTGAG CAAATATACA GGGGGCCGAG GCGTCAGCGT GGGCCCCATC 720  
 CTCAGCAGCA GTGCCTCGGA TATCTTCTGC GACAATGAGA ATGGGCTAA CTTCCTTTTC 780  
 CACAACCGGG GCGATGGCAC CTTTGTGGAC GCTGCGGCCA GTGCTGTGTG GAGCGACCCC 840  
 CACGAGCATG GCGGAGGTGT CGCCCTGGCT GACTTCAACC GTGATG3CAA AGTGACATC 900  
 GTCTATGCGA ACTGGAATGG CCCCCACCGC CTCTATCTGC AAATGAGCAC CCATGGGAAG 960  
 GTCCGCTTCC TGTCGCTCTT CCGGGGCAAT CAGGGCTTCA ACAACAACCT GCTGCGAGTG 1020  
 ATCAGCGCGG ACTTTGACAA TGACCAAGAG CTGAGATCT TCTTCAACAA CATTCCTTAC 1080  
 CGCAGCTCCT CAGCCCAACG CTTCTTCCG GTCATCCGTA GAGAGCACGG AGACCCCTCT 1140  
 ATCGAGGAGC TCAATCCCGG CGAAGCCTTG GAGCCTGAGG GCGGGGGCAC AGGGGGTGTG 1200  
 GTGACGACT TCGAGGAGA CGGGATGCTG GACCTCATCT TGTCCCATGG AGATGCCATG 1260  
 GCTCAGCGC TGTCGCTCTT CCGGGGCAAT CAGGGCTTCA ACAACAACCT GCTGCGAGTG 1320  
 GTGCCACGCA CCGGTTTGGG GGCCTTTGCC AGGGGAGCTA AGGTGCTGCT CTACACCAAG 1380  
 AAGAGTGGGG CCCACCTGAG GATCATCGAC GGGGCTCAG GCTACCTGTG TGAGATGGAG 1440  
 CCGTGGGCAC ACTTTGGCCT GGGGAAGGAT GAAGCCAGCA GTGTGGAGGT GACGTGGCCA 1500  
 GATGGCAAGA TGTGAGCCG GAACGTGGCC AGCGGGGAGA TGAACCTCAGT GCTGGAGATC 1560  
 CTCTACCCCC GGGATGAGGA CACACTTCAG GACCCAGCCC CACTGGAGTG TGGCCAAGGA 1620  
 TTCTCCAGC AGGAAATGG CCATTGATG GACACCAATG AATGCATCCA GTTCCCATTC 1680  
 GTGTGCCCTC GAGACAAGCC CGTATGTGTC AACACCTATG GAAGCTACAG GTGCCGGAAC 1740  
 AACAGAAGAT GCACTCGGGG CTACGAGGCC AAGGAGGAT GCACAGCCTG GGTGGGGACT 1800  
 CTGGGCCAGT CACGGGGCCC CGGCCCCACC ACCCCACCGC CTGCTGCTGC CACTGCCGCT 1860  
 GCTGCTGCCG CTGCTGGAGC TGCCACTGCT GCACCGGTCC TCGTAGATGG AGATCTCAAT 1920  
 CTGGGGTGGG TGGTTAAGGA GAGCTGCGAG CCCAGCTGCT GAGCAGGGGT GGGACATGAA 1980  
 CCAGCGGATG GAGTCCAGCA GGGGAGTGGG AAAGTGGGCT TGTGCTGCTG CCTAGACAGT 2040  
 AGGGATGTAA AGGCCTGGGA GCTAGACCTT CCCAAGCCC ATCCATGCAC ATTACTTAGC 2100  
 TAACAATTAG GGAGACTCGT AAGGCCAGGC CCGTGTCTGG GCACATAGCT GTGATCAG 2160  
 CAGACAGGCT CGCTGCCCTG ATGGCGCTTA CATTCAGTG GGTCTAATGA CCATATCTTA 2220  
 GGACACAGAT GTGCCAGGG AGGTGGTGTG ACTGCACAGG AAGTATGAGG ACTTTAGTGT 2280  
 CCTGAGTTCA AATCCTGATT CAGGAACCTA CAAAGCTATG TGACCTTACA CCAGTCACTT 2340  
 AACTTGTAG CCATCCATTA TCGCATCTGC AAAATGGGGA TTAAGAATAG AATCTTGGGG 2400  
 TTAGTGTGGA GATTAGATTA AATGTATGTA AGACACTTGG CACAAACCT GGCACATAGT 2460  
 AAAGGCTCAA TAAAACAAAG TGCCTCTCAC TGGGCTTTGT CAACAG

Seq ID NO: 471 Protein sequence  
 Protein Accession #: CAC08451

1 11 21 31 41 51  
 MSRLPFLLL LWFLPITEGS QRAEPMPTAV TNSVLPDPYD SNPTQLNYGV AVTDVHDGDD 60  
 FEIVVAGYNG FNLVLKYDRA QKRLVNIADV ERSSPYALR DRQGNALGVT ACDIDGDGRE 120  
 EIYFLNTNNA FSGVATYTDK LPIKFRNRWE DILSDEVNVA RGVASLPAGR SVACVDRKGS 180

GRYSIIYANY AYGNVGPDAL IEMDPEASDL SRGILALRDV AAEAGVSKYT GGRGVSVGPI 240  
 LSSASDIFC DNENGNPNFLF HNRGDGTFVD AAASAGVDDP HQHGRGVALA DFNDRGKVDI 300  
 VYGMWNGPHR LYLQMSHKGK VRFRDIASPK FSPMSPVRTV ITADFDNDQE LEIPFNMIAY 360  
 RSSSANRLFR VIRREHGDPL IEELNPGDAL EPEGRGTGGV VTFDFDGMGL DLILSHGESM 420  
 AQLPLSVFRGN OGFNNWNLVR VPRTRFGAFA RGAKVVLVTK KSGAHLRIID GSGSYLCEMS 480  
 PVAHFLGKLD EASSVEVWTF DGKMSVRNVA SGMNNSVLBI LYPRDEDTLQ DPAPLECGQG 540  
 FSQENGHCMT DTNCEIQFPF VCPDRKPCV NTYGSYRCRT NKKCSRGYEP NEDGTACVGT 600  
 LGQSPGPRPT TPTAAATAA AAAAAGAATA APVLVDGDLN LGSVVKESCE PSC

Seq ID NO: 472 DNA sequence  
 Nucleic Acid Accession #: FGENESH  
 Coding sequence: 1..4794

1 11 21 31 41 51  
 15 ATGGCGTGTG CGGGAGGACT CCCAGCCCGT TGCTCTGGTT GGATGGGACT GGGTGGGCCC 60  
 AGCGGCTCCT CCCAGCATC CCTCCCAT TCCTCCTCCA GGTACATGG ACCCAACCTG 120  
 GTTCGTAAGT ATGACCGGCG CCAGAAGCGG CTGGTGAACA TCGCGGTGCA TGAGCGCAGC 180  
 20 TCACCTTACT ACGCGCTGGG GGACCGGCAG GGAACGCCA TCGGGGTGAC AGCCTGCGAC 240  
 ATCGACGGGG ACGCGCGGGA GGAGATCTAC TTCTCTCAACA CCAATATATG CTCTCGGGC 300  
 CACAGCAGCT CAGCGCAGGT CCTCTCTGGG CTCCACAGAA ACAGGCTGTG GCTGAAGCCT 360  
 CCACCTACAA CCCCTGCAGG CCTCCTGGGT CTGCCTCCAC TCAGCGGAAG GGACTTTTCC 420  
 TCCTCCCTGG GTACAGCTTC TCGGACAGC AGGCAGGGAG AGAGGGTGCC GGTTCCTTCG 480  
 25 TGTCGGGGTG GACTGAGACC TACCATGAA CCAGAACCAT TTCTCTGAG ACCCAATATC 540  
 GGGGTGGCCA CGTACACCGA CAAGTTGTTC AAGTTCGCA ATAACCGGTG GGAAGACATC 600  
 CTGAGCGATG AGGTCAACGT GGCCCGTGGT GTGGCCAGCC TCTTTGCCGG ACGCTCTGTG 660  
 GCCTGTGTGG ACAGAAAGGG CTCTGGACGC TACTCTATCT ACATTGCCAA TTACGCTTAC 720  
 GGTAAATGGG GCCCTGATGC CCTCATGAA ATGGACCTG AGGCCAGTGA CCTCTCCCGG 780  
 30 GGCATTCTGG CGCTCAGAGA TGTGGCTGCT GAGGCTGGGG TCAGCAATA TACAGAGGCG 840  
 TTCTCCACCA CTGCTCTCC AAGCATTGGT GAGATATCTG GCAGAACCGA GGAGCGGGAA 900  
 GGAGGAGACC CAGAGCAGCT AGATGAGGAG CACAGTGGGG ATGGAAGCAC CAGCCAACTG 960  
 TGCCGGCTGG GCTGGAAGGA CGGGCAGTTC AAGGAAGAAG CAGCAGCTTT GGTGGAGGAA 1020  
 CAGAGGGAGG CTGGGGCAGC TGGCGTGCCC AGAGGACGTG TCGAACAGC TCTGCAGACT 1080  
 35 TCCAAAGACC ATTTGGCTGA CAAGAACCTA TTTGGCCCAT CATGTACTA TTCTGTCTGC 1140  
 GCGCTTCTC CAGCCACCCC TTTCCCTGCC CGCCAGCCCC CCCAACACTA CCTGTAGACC 1200  
 CCCCTGTGCA CTGAGCTAAT GACACATGGA CGTCTGGCTG GAAACTAGC CCGGAGTGTG 1260  
 CCCACCCCC GAGCCCCAGG AATGGACCCC AAATGTAAGG GCGCCATGC TGAGCCCCGC 1320  
 CTGATGGCTG AGGCTTTGGG CGCGTGGCCA CGCTCAGCA CCACTGTGTG GCCAGGGGGC 1380  
 40 CTGAGAAAGT GGGAGGAAAG CAGGCAGAA GGGCAGGCCA TGTCAGATG TGCACTCAGG 1440  
 GAGCTGGGAG GTCCCTGGAG CCAAGCCACA CAGCAGCTGC CTGCTAGAGA GCTGTATGAC 1500  
 CTGGAGAAAC CTCCTATTTT ACAAGAAACA GACGAGATC CAGGAGGAG AAGGGACTCG 1560  
 CCCAAGTCA CACAGGAGTG CCATCTAGTG GCCACCATGC CAGCTCTCGG GGGACTCGAG 1620  
 GGCCCGGGA GGTGGCCAA GCGAGAGATT GGGAGAGAGA CTGGGGCAGT AGGAAGACCA 1680  
 45 CTCTCCATC CCTGTGCTCC CACTTCCCC AGCTGCTTGA GGCCTCTTGA AGCCGGGACA 1740  
 GTGCGGGGAG CTGCGCTGCC TGGGAATCCT GGGAACTGGG TTCTGGACAT GGCCAGGGCC 1800  
 CTGGCGTGGA ACCAGATGGA AAAAGAGGAG GGGAAAGATT ATGGAGACCA TGAGCCGACA 1860  
 TTTAGGCTCA GGAAGCAGC GGAAGCAGAA TTCCCCCAG GCTCCTCTGA GGAGCCTCTG 1920  
 CTGAGTTTCC CTGCTAGCTC CAGAGGCAGC CCGTCTCTCC AGGTGGGCGT GGGGCTTGCT 1980  
 50 TCTGCCACTC ACTGTGGGTC GATGTCTTTT CTAGGGGGCC GAGGCGTCAG CGTGGGCCCC 2040  
 ATCCTCAGCA GCAGTGCTTC GGATATCTTC TGCGACAATG AGAATGGGCC TAACCTCTCTT 2100  
 TTCCACAACC GGGCGATGGG CACCTTTGTG GACGCTGCGG CCACTGTCTG ACGTCTTTTA 2160  
 GCCTTCATCG TTACCTCAA ATATCACCTC TGCAAGATTT TTCTCACTC CCGTGTCCAC 2220  
 CTAGCAGAAA CTGTCTCTGC CTCTCTCTGC TGCCCGTGGC ATGCAGGTCT TCTTCAGGCT 2280  
 CCACATGTCC ATCATGTTT GTCTATGAGC TTTACAAGGA CCGGCTCAGG GTTCTATTCA 2340  
 55 TTCTTGACGC AAGGCTTGGC CTCCAGTGCC CACCGAGGA CACTCAGCCT CCAGGGTTCT 2400  
 CAGGGGGCCC CACCTGCTCT TCTGGCAAGA GCTCCCTGTG TCCTGGGGTC TCTGATCCCC 2460  
 ACTGCTTATT ACATTGTCTC GTGTCTGTCC ATCCACAGGA GCCTGTATGAC CCACAGCTAT 2520  
 TTGTCTCTCG AAAGAGTCAA CGTGGGTGTG GACGACCCCC ACCAGCATGG GCGAGGTGTC 2580  
 GCGCTGGCTG ACTCAACCG TGATGGCAAA GTGGACATCG TCTATGGCAA CTGGAATGGC 2640  
 60 CCCACCGCC TCTATCTGCA AATGAGCACC CATGGGAAGG TCCGCTTCCG GGACATGGCC 2700  
 TCACCCAAGT TCTCCATGCC CTCCTCTGTC CGCACGGTCA TCACCGCCGA CTTTGACAAT 2760  
 GACCAGGAGC TGGAGATCTT CTTCACAAC ATTGCCTACC GCAGCTCCTC AGCCAAACCGC 2820  
 CTCTTCGAT GCTCATCTCT GGCTGTGTGC TCTTCATCT TGACAGCTGG TGGGAGGAAC 2880  
 65 GGTGAGGAG AAGGTTTAAG AATCAGAAGG GAGGGGTTCC CAGGGCCAGG GGGTCAGGCC 2940  
 AAGGTCAACA CAGTCCCTCT GATGAAGAAA CAGAAAGGAA GGAAGGACGA GGACTGGGCA 3000  
 AGAGGCTGTG GGAATGCAGG GCAAAGCCTG GCCAAGGAGC CGGCTCTGTC TATTGCAGGG 3060  
 AAAGGGAAGG GAAATGTGGC CCAAAGTGTG CCCAGAACCC AAGCGCCACA AGATACAAAG 3120  
 CCACACTACC ACAAAAGGGG GCTACAGGGT CCAATCACTA CCAGGAAAGG GGGCTACGGG 3180  
 70 GTCCAACTAC TACCAGGAAA AGGGGCTACG GGGTCCAATC ACTACCAGGA AAAGGGGCTA 3240  
 CGGGGTCCAA TCACTACCA GAAAAGGGGC TACGGGGTCC AATCACTACC AGGAAAAGGG 3300  
 GCTACGGGCT CCAATCACTA CCAGGAAAGG GGGCTACAGG GTCCAATCAC TACCAGGAAA 3360  
 AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA CAGGGTCCAA TCACTACCAC 3420  
 75 AGAAAAGGGC TACGGGCTCC AATCACTACC AGGAAAAGGG GCTACGGGGT CCAATCACTA 3480  
 CCAGGAAAGG GGGCTACAGG GTCCAATCAC TACCAGGAAA AGGGGCTACG GGGTCCAATC 3540  
 ACTACCAGGA AAAGGGGCTA CGGGCTCCAA TCACTACCA GAAAAGGGGC TACGGGGTCC 3600  
 AATCACTACC AGGAAAAGGG GCTACAGGGT CCAATCACTA CCAGGAAAGG GGGCTACAGG 3660  
 GTCCAATCAC TACCAGGAAA AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA 3720  
 80 CGGGGTCCAA TCACTACCAG GAAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGAG 3780  
 GCTATGGGGT CCAATCACTA CCAGGAAAGG GGGCTACGGG GTCCAATCAC TACCAGGAAA 3840  
 AGGGGCTATG GGTGTCAAATC ACTACCAGG AAAAGGGGCTA CGGGGTCCAA CGTATCCGT 3900  
 AGAGAGCAGC GAGACCCCTT CATCGAGGAG CTCAATCCCG GCGACGCTTT GGAGCCTGAG 3960  
 GGCAGGGGCA CAGGGGTGTG GGTGACCGAC TTCGACGGAG ACGGATGTCT GGACCTCATC 4020  
 85 TTGTCCCATG GAGAGTCCAT GGCTCAGCGC CTGTCCGTCT TCCGGGGCAA TCAGGGCTTC 4080  
 AACCAACACT GCTTCGAGGT GGTGCCACGC ACCCGGTTTG GGGCCTTTGC CAGGGGAGCT 4140  
 AAGTGTGTGC TCTACCCCAA GAAGAGTGGG GCCACCTGCA GGATCATCGA CGGGGGCTCA 4200  
 GGCTACCTGT GTGAGATGGA GCCCGTGGCA CACTTTGGCC TGGGGGAAGGA TGAAGCCAGC 4260  
 AGTGTGGAGG TGACGTGGCC AGATGGCAAG ATGGTGAGCC GGAACGTGGC CAGCGGGGAG 4320

WO 02/086443

PCT/US02/12476

ATGAAGCTAG TGCTGGAGAT CCTCTACCCC CGGATGAGG ACACACTTCA GGACCCAGCC 4380  
CCACTGGAGT GTGGCCCAAGG ATTTCTCCAG CAGGAAAATG GCCATTGCGT GGACACCAAT 4440  
GAATGCATCC AGTTCCCATTT CGTGTGCCCT CGAGACAAGC CCGTATGTGT CAACACCTAT 4500  
GGAAAGTACA GGTGCGCGAC CAACAAGAAG TGCAAGTCGG GCTACGAGCC CAACGAGGAT 4560  
GGCAGACGCT GCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG 4620  
CCCAAAAAGG AGCTGCAACT TTCCCAAGGC ATCTGCACCC CCGTCTGGTC CTTTTCCTG 4680  
CCGGGTGACC GGTGCTCCT CAAAAGAGCT CAGCTCCAGG CTGCTCCAG CACCCTTCTC 4740  
CAGAAAGCTC CAGGTATTC AGAAGCCCA GTGTATGAAC AAGATCAGGA ATAA

Seq ID NO: 473 Protein sequence  
Protein Accession #: F0ENESH predicted

1 11 21 31 41 51  
MACPGGLPAR CSGNMGLGGP SGSSPASPPH SSSRYNGPNL VLKYDRAQKR LVNIAVDERS 60  
SPYYALRDRQ CNAIGVTACD IDGDGRIEYI FLNTNNAFSG HSSSAQVPSG LHRNRPVLKP 120  
PPTTFAGLLG LPLPSGRDFS SSLGQASPDG RQGERVFPVC CRGGLRPTHG PEPFLLRPKS 180  
GVATYTDKLF KPRNRRWEDI LSDEVNVARG VASLFAGRSV ACVDRKGSGR YSIYIANYAY 240  
GNVGPDALIE MDPEASDLSE GILALRDVAA EAGVSKYTEG FSHTASPSIG EISGRTEERE 300  
GGDPPEADEE HSGDGSSTQL CRLGWKDGQF KERAALVEE QREAGAAGVP RGRVRTALQT 360  
SKSHLADKNL FGPPCYYSVC APSPAHPPPA RQAPQHYVPA PLVTQLMTHG RLAKGLARSV 420  
PHPRAPGMDP KCKGRHAEPG LMAEALGAMP ALSTTVVPGG LRSNEESRQK GQAMSRCLAR 480  
ELGGPWSQAT QHLFARELYD LGEPPIQRT DGDPRRRDS PKVTQECHLV ATMPALGGLE 540  
QFGRVAKREI GRBTGAVGRF LSHPLVFNFF SCLRPLEAGT VPGAALPGNP GNVVLDMAKA 600  
LAWNQMEKEE GKINGDHEPR FRLRKAREAE FPPGSSEELP LQPPSGLRGS PVLQVGLGLA 660  
SATHCGSMF LGGRGVSVGP ILSSASDIF CDNENGPNFL FHNRGDGFV DAAASAERRL 720  
AFIVELKXHL CRDPFSLCH LAETGPSSSC CPWHARLLQA PHCHGLSMS FTRTGRSPYS 780  
FLTQGLASSA HRRTLSLQGS QGAPPCLLAR APCVLGSLIP TAYYIVLWSA IPESLMTSHY 840  
LSSERVNVGV DDPHQHGRGV ALADFNDRGK VDIVYGNWNG PHRLYLQMSH HGKVRFRDIA 900  
SPKFSMPSPV RTVITADFEN DQELRIFPNM IAYRSSANR LPRCSILARG SSSLTAGGRN 960  
GQGEGLRIRR GGFPGPGGQA KVNTPGLMKK QKGRKDEDNA RGCGNAGQSL AKEPASAIAG 1020  
KKGKNAQSV PRTQAPQDTK PHYHKKGLQG PITTRKRGYV VQSLPGKQAT GSNHYQEKGL 1080  
RGPIITRKRG YGVQSLPGKG ATGSNHYQEK GLQGPITTRK RGYGLQSLPG KGATGSNHYH 1140  
RKGLRAPITT RKRGYGVQSL PGKATGSNH YQEKGLRGP I TTRKRGYGLQ SLPGKATGS 1200  
NHYSKGLQG PITTRKRGYV VQSLPGKQAT GSNHYQEKGL RGPITTRKRG YGLQSLPGKE 1260  
AMGSNHYQEK GLRAPITTRK RGYGVQSLPG KGATGSNVIR REHGDPLIEE LNPDALEPE 1320  
GRGTGVVTD FLDGDMLDLI LSHGESMAQP LSVFRGNQGF NNNWLVVPR TRFAGFARGA 1380  
KVVLVTKKSG AHLRIIDGGS GYLCEMEPVA HFGLGKDEAS SVEVTWPDGK MVRNVSASGE 1440  
MNSVLILYP RDEDTLQDPA PLSCGGGFSQ QENGHCMNDT ECIQFPFVCP RDKFVVCNTY 1500  
GYSYRCITNK CSRGYEPNED GTACVGTGLG SRHTMTWKFR PKKSLQLSQG ICTPVWSPFL 1560  
PGCRLLKKRA QLQAAPSTLL QKAPGIPBAQ VYEQDQE

Seq ID NO: 474 DNA sequence  
Nucleic Acid Accession #: NM\_003661.1  
Coding sequence: 1..1152

1 11 21 31 41 51  
ATGAGTGCAC TTTTCCTTGG TGTGGGAGTG AGGCGAGAGG AAGCTGGAGC GAGGGTGCAA 60  
CAAAAGCTTC AGTGGGGAC AGATACTGGA GATCCTCAA GATAGCCCT CGGTGACTGG 120  
GCTGCTGGCA CCATGGACCC AGAGAGCAGT ATCTTTATTG AGGATGCCAT TAAGTATTTT 180  
AAGGAAAAGG TGAGCACACA GAATCTGCTA CTCTGCTGA CTGATAATGA GGCTGGAAC 240  
GGATTCTGTG CTCTGCTGA ACTGCCAGG AATGAGGCG AGTGGCTCG TAAAGCTCTG 300  
GACAACTCTG CAAGACAAAT GATCATGAAA GACAAAATCT GGCACGATAA AGCCAGCAG 360  
TACAGAACT GGTTCCTGAA AGAGTTCTCT CGGTGAAA GTGAGCTTGA GGATAACATA 420  
AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGTCC ACAAGGCAC CACCATCGCC 480  
AATGTGGTGT CTGGCTCTCT CAGCATTTC TCTGGCATCC TGACCTCGT CGGCATGGGT 540  
CTGGCACCTT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA 600  
ATCACAGCCC CTGTGACCGG GATTACACAG AGTACCATGG ACTACGAAA GAAGTGCTGG 660  
ACACAAGCCC YAGCCACGCA CTGTGTCATC AAAAGCCTTG ACAAAATTGA GGAGGTGAGG 720  
GAGTTTTTGG GTGAGAACAT ATCCAACITT CTTCCTTAG CTGGCAATAC TTACCACTC 780  
ACACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA 840  
GTACCGCATG CCTCAGCCTC AOGCCCCCGG GTCACTGAGC CAATCTCAGC TGAAAGCGGT 900  
GAACAGGTGG AGAGGGTTAA TGAACCCAGC ATCCTGGAAT TGAGCAGAGG AGTCAAGCTC 960  
ACCGATGTGG CCCCTGTAA CTTCTTCTT GTGCTGGATG TAGTCTACCT CGTGTACGAA 1020  
TCAAGCACT TACATGAGG GGCAGAGTCA GAGACAGCTG AGGAGCTGAA GAAGGTGCT 1080  
CAGGAGCTGG AGGAGAAGCT AACATTCTC AACAAATATT ATAAGATTCT GCAGGCGGAC 1140  
CAAGAACTGT GA

Seq ID NO: 475 Protein sequence  
Protein Accession #: NP\_003652.1

1 11 21 31 41 51  
MSALPLGVGV RAEAGARVQ QNVPSGTDG DPQSKPLGDW AAGTMDPESS IFIEDAIKYF 60  
KERVSTONLL LLLTDNEAWN GFVAAELPR NEADELRKAL DNLARQMIMK DKNWHDKQOQ 120  
YRWFLKEFP RLKSELEONI RRLRALDGV QKVEKGTIA NVVSGSLSTIS SGILTLVGMG 180  
LAPFTEGSSL VLLEPGMELC ITAALTGITS STMDYGGKWW TQAQAEHLVI KSLDKLKEVR 240  
EPLGENISNF LSLAGNTYQL TRGIGKDIRA LRRARANLQS VPHASASRPR VTEPISAESG 300  
EQVERVNEPS ILEMSRGVKL TDVAPVSPFL VLDVVVLVYE SKHLHEGAKS ETAEELKKVA 360  
QELEEKLNL NNNYKILQAD QEL

Seq ID NO: 476 DNA sequence  
Nucleic Acid Accession #: NM\_014452.1  
Coding sequence: 1..1968

1 11 21 31 41 51



	ATGGGGACCT	CTCCGAGCAG	CAGCACCGCC	CTCGCCTCCT	GCAGCCGCAT	GGCCCGCCGA	60
	GCCACAGCCA	CGATGATCGC	GGGCTCCCTT	CTCCTGCTTG	GATTCCCTTAG	CACCACCACA	120
5	GCTCAGCCAG	AACAGAAGGC	CTCGAATCTC	ATTGGCACAT	ACCGCCATGT	TGACCGTGCC	180
	ACCGGCCAGG	TGCTAACTTG	TGACAAGTGT	CCAGCAGGAA	CCTATGTCTC	TGAGCATTTG	240
	ACCAACACAA	GCTTGCAGGT	CTGCAGCAGT	TGCCCTGTGG	GGACCTTTAC	CAGGCATGAG	300
	AATGGCATAG	AGAAATGCCA	TGACTGTAGT	CAGCCATGCC	CATGGCCAA	GATTGAGAAA	360
	TTACTTTGTG	CTGCTTGTG	TGACCGAGAA	TGCACTTGCC	CACCTGGCAT	GTTCCAGTCT	420
10	AAOGCTACCT	GTGCCCCCA	TACGGTGTGT	CCTGTGGGT	GGGGTGTGCG	GAAGAAAGGG	480
	ACAGAGACTG	AGGATGTGCG	GTGTAAAGCAG	TGTGCTCGGG	GTACCTTCTC	AGATGTGCCT	540
	TCTAGTGTGA	TGAATATCAA	AGCATACACA	GACTGTCTGA	GTGAGAACCT	GSTGGTGATC	600
	AAGCCGGGGA	CCAAGGAGAC	AGACAACGTC	TGTGGCACAC	TCCGTCCTT	CTCCAGCTCC	660
	ACCTCAGCTT	CCCCTGGCAC	AGCCATCTTT	CCAGCCCTTG	AGCACATGGA	AAOCCATGAA	720
15	GTCCCTTCTT	CCACTTATGT	TCCCAAAGGC	ATGAACCTCA	CAGAATCCAA	CTCTTCTGCC	780
	TCTGTTAGAC	CAAAGGTACT	GAGTAGCATC	CAGGAAGGGA	CAGTCCCTGA	CAACACAAGC	840
	TCAGCAAGGG	ATCCAGTATG	CGTGAAACAG	ACCCCTCCAA	ACCTTCAGGT	AGTCAACAC	900
	CAGCAAGGCC	CCCAACACAG	ACACATCCTG	AAGCTGTCTG	CGTCCATGGA	GGCCACTGGG	960
	GGCGAGAGT	CCAGCAGCGC	CATCAAGGGC	CCCAAGAGGG	GACATCTAG	ACAGAACCTA	1020
20	CACAAGCATT	TTGACATCAA	TGAGCATTG	CCCTGGATGA	TTGTGCTTTT	CCTGCTGCTG	1080
	GTGCTTGTGG	TGATTGTGGT	GTGCAGTATC	CGSAAAAGCT	CGAGGACTCT	GAAGAAAGGG	1140
	CTCCGGCAGG	ATCCAGTATG	CATTGTGGAA	AAGGCAGGGC	TGAAGAAATC	CATGACTCCA	1200
	ACCCAGAAC	GGGAGAAATG	GATCTACTAC	TGCAATGGCC	ATGGTATCGA	TATCCTGAAG	1260
	CTTGTAGCAG	CCCAAGTGGG	AAGCCAGTGG	AAAGATATCT	ATCAGTTTCT	TTGCAATGCC	1320
25	AGTGAGAGGG	AGGTTGTCTG	TTTCTCCAAT	GGGTACACAG	CCGACCACGA	CGGGGCTTAC	1380
	GCAGCTCTGC	AGCACTGGAC	CATCCGGGGC	CCCAGAGCCA	GCCTCGCCCA	GCTAATTAGC	1440
	GGCTCGCGCC	AGCACCGGAG	AAACGATGTT	GTGGAGAAGA	TTGCTGGGCT	GATGGAAGAC	1500
	ACCAACCCAG	TGGAATCTGA	CAAACTAGCT	CTCCCGATGA	GCCCGAGCCC	GCTTAGCCCG	1560
	AGCCCATCC	CAGGCCCAAA	CGCGAAACTT	GAGAATTCCG	CTCTCCTGAC	GGTGGAGCCT	1620
30	TCCCCACAG	ACAAGAACAA	GGGCTTCTC	GTGGATGAGT	CGGAGCCCTT	TCTCGCTGT	1680
	GACTCTACAT	CAGCGGCTC	CTCCGCGCTG	AGCAGGAACG	GTTCCTTTAT	TACCAAGAA	1740
	AAGAAGGACA	CAGTGTTCGG	GCAGGTACGC	CTGGACCCCT	GTGACTTGCA	GCCTATCTTT	1800
	GATGACATCG	TCCACTTTCT	AAATCCTGAG	GAGCTGCGGG	TGATTGAAGA	GATTCCCCAG	1860
35	GCTGAGGACA	AAGTAGACCG	GCTATTGGA	ATTATTGGAG	TCAAGAGCCA	GGAAGCCAGC	1920
	CAGACCTCC	TGACTCTGT	TTATAGCCAT	CTTCTGACC	TGCTGTAG		

Seq ID NO: 477 Protein sequence  
Protein Accession #: NP\_055267.1

	1	11	21	31	41	51	
40	MGTSPPSSSTA	LASCSRIARR	ATATMIAGSL	LLLGFLSTTT	AQPEQKASNL	IGTYRHVDRA	60
	TGQVLTCDKC	PAGTYVSEHC	TNTSLRVCS	CPVGTFRHE	NGIEKCHDCS	QPCPWPMEK	120
	LPCALITDRE	CTCPGMFQJS	NATCAPHTVC	PVGWGVKKGS	TETEDVRCKQ	CARGTFSDVP	180
45	SSVMKCKAYT	DCLSQLNVI	KPGTKETDNV	CGTLPSPSSS	TSPSPGTAIF	PRPEHMETHE	240
	VPSSTVVPKG	MNSTESNSGA	SVRPKVLSSI	QEGTVPDNTS	SARGKEDVNK	TLPNLQVNVH	300
	QGGPHRHIL	KLPSMEATG	GEKSSTFIKG	PKRGHPRQNL	HKHFDINEHL	PWMIVLFLLL	360
	VLVVIVVCSI	RKSSRTLKKG	PRQDPSAIVE	KAGLKKSMTP	TQNRKWIYY	CNGHGIDILK	420
	LVAQVGSQW	KDIYQFLCNA	SEREVAAPSN	GYTADHERAY	AALQHWITRG	PEASLAQLIS	480
50	ALRQHRNRDV	VEKIRGLMED	TTQLETDKLA	LPMSPSPLSP	SPIPSFNAKL	ENSALLTVEP	540
	SPQDKNKGFF	VDSEPELLRC	DSTSSGSSAL	SRNGSFITKE	KKDTVLQRVR	LDPDCLQPIF	600
	DDMLHFLNPS	ELRVIEBIPQ	AEDRLDLFE	IIGVKSQEAS	QTLDSVYSH	LPDLL	

Seq ID NO: 478 DNA sequence  
Nucleic Acid Accession #: XM\_044533  
Coding sequence: 238..2751

	1	11	21	31	41	51	
60	GCTCTGCCCA	AGCCGAGGCT	GGGGGGCGGG	CGCGGGCGGG	AGGACTGCGG	TGCCCGCGGG	60
	AGGGGCTGAG	TTTGCCAGGG	CCCACTTGAC	CCTGTTTCCC	ACCTCCCGCC	CCCCAGGTCC	120
	GGAGGCGGGG	GCCCCCGGGG	CGACTCGGGG	CGCGACCGCG	GGCGGAGCT	GCCGCCCGTG	180
	AGTCCGGCCG	AGCCACCTGA	GCCCGAGCCG	CGGGACACCG	TGCTCTCTGC	TCTCCGAATG	240
	CTGCGCACCG	CGATGGGCTT	GAGGAGCTGG	CTCGCCCGCC	CATGGGCGCG	GCTGCCCGCT	300
65	CGGCCACCGC	TGCTGTGCTT	CCTGCTGCTG	CTGCTCTCTG	TGCAGCGCGC	GCCTCGAGCC	360
	TGGGCGCTCA	GCCCCCGGAT	CAGCCTGCTT	CTGGGCTCTG	AAGAGCGGCC	ATTCTCTAGA	420
	TTGGAAGCTG	AACACATCTC	CAACTACACA	GCCCTTCTGC	TGAGCAGGGA	TGGCAGGACC	480
	CTGTACGTGG	GTGCTCGAGA	GGCCCTCTTT	GCACTCAGTA	GCAACCTCAG	CTTCTTGCCA	540
	GGCGGGGAGT	ACCAGGAGCT	GCTTTGGGGT	GCAGACGCGA	AGAAGAAACA	GCAGTGCAGC	600
70	TTCAAGGGCA	AGGACCCACA	CGCGGACTGT	CAAACTACA	TCAAGATCCT	CCTGCCCGCTC	660
	AGCGGCAGTC	ACCTGTTTAC	CTGTGGCACA	GCAGCCTTCA	GCCCCATGTG	TACCTACATC	720
	AACATGGAGA	ACTTACCCCT	GGCAAGGGAC	GAGAAGGGGA	ATGTCCTCTC	GGAAGATGGC	780
	AAGGGCCGTT	GTCCCTTCGA	CCCGAATTTT	AAGTCCACTG	CCCTGCTGGT	TGATGGCGAG	840
	CTCTACACTG	GAACAGTCAG	CAGCTTCCAA	GGGAATGACC	CGGCCATCTC	GCGGAGCCAA	900
75	AGCCTTGCGC	CCACCAAGAC	CGAGAGCTCC	CTCAACTGGC	TGCAAGACCC	AGCTTTTGTG	960
	GCCTCAGCCT	ACATTCTCTG	GAGCCTGGGC	AGCTTGCAAG	GCAGATGATG	CAAGATCTAC	1020
	TTTTTCTTCA	GCGAGACTGG	CCAGGAATTT	GAGTTCTTTG	AGAACACCAT	TGTGTCCCCT	1080
	ATTGCCCGCA	TCTGCAAGGG	CGATGAGGGT	GGAGAGCGGG	TGCTACAGCA	GCGCTGAGCC	1140
	TCCTTCTCTA	AGGCCAGCT	GCTGTGCTCA	CGGCCCGAGC	ATGGCTTCCC	CTTCAACGCT	1200
80	CTGCAAGATG	TCTTACGCTT	GAGCCCCAGC	CCCCAGGACT	GGCGTACACC	CTTTTCTTAT	1260
	GGGGTCTTCA	CTTCCAGTGG	GCACAGGGGA	ACTACAGAA	GCTCTGCCGT	CTGTGTCTTC	1320
	ACAATGAAGG	ATGTGCAGAG	AGTCTTCAGC	GGCTTCTACA	AGGAGGTGAA	CCGTGAGACA	1380
	CAGCATGTGT	ACACGCTGAC	CCACCGGCTG	CCACACCC	GGCCTGGAGC	GTGCTATCAC	1440
	AACAGTGCCC	GGGAAAGGAA	GATCAACTCA	TCCCTGACAG	TCCAGAGCCG	CGTGCTGAAC	1500
	TTCCCTAAGG	ACCACTTCTC	GATGGACGGG	CAGGTCCGAA	GCAGCATGCT	GCTGCTGAG	1560
85	CCCCAGGCTC	GCTACAGCGG	CATGGCTGTA	CACCGGCTCC	CTGGCTTGCA	CCACACTTAC	1620
	GATGTCTCTT	TCTTGGGCAC	TGTTGACGGC	CGGCTCCACA	AGGCAGTGAG	CATGGGCCCC	1680
	CGGGTGACCA	TCATTGAGGA	GCTGCAGATC	TTCTCATCGG	GACAGCCGCT	GCAGAACTCG	1740



CTCCTGGACA CCCACAGGGG GCTGCTGTAT GGGGCTCAC ACTCGGGCGT AGTCCAGGTG 1800  
CCCATGGCCA ACTGACGCTT GTACAGGAGC TGTGGGACT GCTCCTCGC CCGGACCCC 1860  
TACTTGTCTT GGAGCGGCTC CAGCTGCAAG CACGTGAGCC TCTACAGCC TCAGCTGGCC 1920  
ACCAGGCGGT GGATCCAGGA CATCGAGGA GCCAGCGCA AGGACCTTTG CAGCGGTCT 1980  
TCGGTGTGT CCGCTCTTT TGTACCAACA GGGGAGAAGC CATGTGAGCA AGTCCAGTTC 2040  
CAGCCCAACA CAGTGAACAC TTTGGCTGTC CCGCTCCTCT CCAACCTGGC GACCCGACTC 2100  
TGGCTACGCA ACGGGGCCCC CGTCAATGCC TCGGCTCCTT GCCAGTGCT ACCCACTGGG 2160  
GAGCTGTGTC TGGTGGGCAC CCAACAGCTG GGGGAGTTC AGTGTGGTC ACTAGAGGAG 2220  
GGCTTCCAGC AGCTGGTAGC CAGCTACTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280  
CAACAGATG AGGGTGGCAG TGTACCGGTC ATTATCAGCA CATCGCGTGT GAGTGACCA 2340  
GCTGTGGCA AGGCCAGCTG GGTGCGAGC AGGTCTTACT GGAAGGAGTT CCTGGTGATG 2400  
TGCACGCTCT TGTGCTGGC CGTGTGCTC CAGTCTTAT TCTTGCTCTA CCGGCACCGG 2460  
AACAGCATGA AAGTCTTCTT GAAGCAGGGG GAATGTGCCA GCGTGACCCC CAAGACCTGC 2520  
CCTGTGGTGC TGGCCCTGGA GACCCGCCA CTCACGCGCC TAGGGCCCCC TAGCACCCCG 2580  
CTCATCACC GAGGGTACCA GTCCCTGTCA GACAGCCCCC CGGGTCCCG AGTCTTCACT 2640  
GAGTCAGAGA AGAGGCCACT CAGCATCCA GACAGCTTCG TGGAGGTATC CCCAGTGTG 2700  
CCCCGGCCCC GGGTCCGCTT TGGCTCGGAG ATCGGTACT CTGTGGTGTG AGAGCTGACT 2760  
TCCAGAGGAC GCTGCCCTGG CTTAGGGGTC TGTGAATGCT CGAGAGAGGT CACTGAGACC 2820  
TCCCTCTGCG TCTGCTCTTC GTGGAACAG ACCGTGGTGC CCGGCCCTTG GGAGCCTTGG 2880  
GGCCAGCTGG CCTGTGCTC TCCAGTCAAG TAGCGAAGCT CTTACACCCC AGACACCCAA 2940  
ACAGCGTGG CCGCAGAGCT CCTGGCCAAA TATGGGGGCC TGCTAGGTT GGTGAACAG 3000  
TGCTCTTAT GTAACCTGAG CCGTCTGTTT AAAAAACAAT TCCAAATGTG AAATAGAAAT 3060  
GAGAGGGAAG AGATAGCATG GCATGCGACA CACAGCGCTG CTCAGTTCA TGGCCTCCCA 3120  
GGGTGCTGG GATGTCATCC AAAGTGGTTG TCTGAGACAG AGTTGGAAC CCTCACCAAC 3180  
TGGCCTCTTC ACCTTCCCA TTAATCCGCT GCCACCGCT GCCTGTCTC ACTGCAGATT 3240  
CAGGACAGC TGGGCTGCTG TGGCTCTGCT CTTGCCAGTC AGCCGAGGAT GTAGTTGTG 3300  
CTGCGTGT CACACACCTC CAGGACAGC AGGGCTAGT TGGCACTGCG GCCCTCACA 3360  
GGTCTGGG TCGGACCCA CTCTGAGCC TTCCAGCT GTATCAGCT GTGGCCACAC 3420  
GAGAGGACAG CGGAGCTCA GGAGGATT CGTGACAATG TACGCTTTC CTCTCAGATT 3480  
CAGGGAAGAG ACTGTGCTC GCCTTCTCC GTTGTGTGCT GAGAACCGT GTGCCCCCTC 3540  
CCACCATATC CAGCTCGCT CCATCTTGA ACTCAACAC GAGGAACCTA CTGACCCCTG 3600  
GTCTCTCCC CAGTCCCGC TTCACTCC ATCCTCACC TTCTCCACT CTAAGGGATA 3660  
TAAACACTGC CCAGCACAGG GCGCTGAAT TTATGTGTT TTTATACATT TTTAATAAG 3720  
ATGACCTTA TGTCTTTT TAATAAGTC TGAAGAATTA CTGTTT

Seq ID NO: 479 Protein sequence  
Protein Accession #: XP\_044533.3

1 11 21 31 41 51  
MLRTAMGLRS WLAAPWALP PRPPLLLLL LLLLLQPPPP TWALSPRISL PLGSEBRPFL 60  
RFEAEHISNY TALLLSRDRG TLYVGAREAL FALSSNLSFL PGGEYQELLW GADAERKQCC 120  
SPFGKDFQD QONYIKILLP LSGSHLPTCG TAAPSPMCTY INMENFTLAR DEKNVLLD 180  
GKGRCPFPDN FKSTALVVDG ELYTGVVSSF QGNDPAISRS QSLRPTKTES SLNWLQDPAF 240  
VASAYIPRESL GSLQGGDDKI YFFSETGQE FEFFENTIVS RIARICKGDE GGERVLQQRW 300  
TSFLKAQILLC SRPDDGFFPN VLQDVFTLSP SPQDWRDTLP YGVFTSQNHR GITEGSAVCV 360  
TFMKDQVRVF SGLYKEVNRE TQWYVTHP VPTPRPGACI TNSARERKIN SSLQLPDRVL 420  
NFLKDRFLMD GQVRSRMLLL QPQARYQVVA VHRVPGHHT YDVLFLGTG GLRHKAVSVG 480  
PRVHIEELQ IPSSGQPVQN LLLDTHRGLL YAASHSGVVQ VPMANCSLYR SCGDCLLARD 540  
PYCAWSGSSC KHVSLYQPLQ ATRPNIQDIE GASAKDLCSA SSVSPSPFVP TGEKPCQVQ 600  
FPQNTVNTLA CPLLSNLATR LMLRNGAPVN ASASCHVLPF GDLVLVGTQO LGFEQCMSLE 660  
EFGQQLVASV CPVEVEDGVA DQTDGEGSVF VIISTSRVSA PAGGKASWGA DRSYWKFLV 720  
MCTLFVLAFL LPVLPVLYRH RNSMKVFLKQ GECAVHPKT CPVVLFPETR PLNGLPPST 780  
PLDRGYQSL SDSPPGSRVF TSEKRPLSI QDSFVEVSPV CPRPRVRLGS BIRDSV

Seq ID NO: 480 DNA sequence  
Nucleic Acid Accession #: NM\_004217.1  
Coding sequence: 58..1092

1 11 21 31 41 51  
GGCCGGGAGA GTAGCAGTGC CTTGGACCCC AGCTCTCCTC CCCCCTTCTC TCTAAGGATG 60  
GCCCAGAAGG AGAACTCCTA CCGCTGGCCC TACGGCCGAC AGACGGCTCC ATCTGGCCTG 120  
AGCACCTGTC CCCAGCGAGT CCTCCGGAAA GAGCCTGTCA CCCCATCTGC ACTTGTCTC 180  
ATGAGCGGCT CCAATGTCCA GCCCAGAGCT GCCCTGGCC AGAAGGTGAT GGAGATAGC 240  
AGTGGGACAC CCGACATCTT AACGCGGCAC TTCACAATG ATGACTTTGA GATTGGGCGT 300  
CCTCTGGGCA AAGGCAAGTT TGGAAACGTG TACTTGGCTC GGGAGAAGAA AAGCCATTTC 360  
ATCGTGGGTC TCAAGGTCTT CTTCAAGTCC CAGATAGAGA AGGAGGGCGT GGAGCATCAG 420  
CTGCGCAGAG AGATCGAAAT CCAGGCCAC CTGCACCATC CCAACATCCT GCGTCTCTAC 480  
AACTATTTTT ATGACCGGAG GAGGATCTAC TTGATTCTAG AGTATGCCCC CCGCGGGGAG 540  
CTCTACAAGG AGCTGCAGAA GAGCTGCACA TTTGAGAGC AGCGAACAGC CACGATCATG 600  
GAGGAGTTGG CAGATGCTCT AATGTACTGC CATGGGAAGA AGGTGATTCA CAGAGACATA 660  
AAGCCAGAAA ATCTGCTCTT AGGGCTCAAG GGAGAGCTGA AGATTGCTGA CTTGCGTGG 720  
TCTGTGATG CGCCCTCCCT GAGGAGGAAG ACAATGTGTG GCACCCCTGGA CTACCTGCCC 780  
CCAGAGATGA TTGAGGGGCG CATGCACAA TGAAGAGTGG ATCTGTGGTG CATTGGAGTG 840  
CTTTGCTATG AGCTGCTGGT GGGGAACCCA CCTTTGAGA GTGCATCACA CAACGAGACC 900  
TATCGCGGCA TCGTCAAGGT GAGCTAAAG TTCCCGCTT CTGTGCCAC GAGAGCCAG 960  
GACCTCATCT CCAAACTGCT CAGGCATAAC CCCTCGGAAC GGCTGCCCCC GGGCCAGTTC 1020  
TCAGCCACCC CTTGGGTCCG GGCCAACTCT CGGAGGGTGC TGCTCCCTC TGCCCTTCAA 1080  
TCTGTGCTCT GATGGTCCCT GTCAATCACT CGGGTGCCTG TGTGTGATG TCTGTGATG 1140  
TATAGGGGAA AGAAGGATC CTAATCTGTT CCCTTATCTG TTTTCTACT CCTCCTTGT 1200  
TTAATAAAGG CTGAAGCTTT TTGT

Seq ID NO: 481 Protein sequence  
Protein Accession #: NP\_004208

1 11 21 31 41 51

MAQKENSYPW PYGRQTAPSG LSTLPQVRVLR KEPVTPSALV LMSRSNVQPT AAPGQXVMEN 60  
SSGTDPILTR HFTIDDFETG RPLGKKGKPGN VYLAREKKSH FIVALKVLFK SQIEKEGVEH 120  
QLRREIEIQA HLHHPNLLRL YNYFYDRRRI YLILEYAPRG ELYKELQKSC TFDEQRTATI 180  
MEELADALMY CHGKKVIHRD IKPENLLGLL KGEKLIADFG WSVHAPSLRR KTMCGTLDYL 240  
PPMEMIEGRMH NEKVDLWCIG VLCEYELLVGN PPFESASHNE TYRRIVKVDL KPPASVPTGA 300  
QDLISKLLRH NPSERLPLAQ VSAHPWVRAN SRRVLPSPAL QSWA

Seq ID NO: 482 DNA sequence  
Nucleic Acid Accession #: AK055663  
Coding sequence: 38..1423

1 11 21 31 41 51  
15 AGAACGGCTT CCGCGGGGAG CTGTGCAGCT CCTATCATG GGGACAATTC ATCTCTTTCG 60  
AAAAACACAA AGATCCTTTT TTGGCAAGTT GTTACGGGAA TTTAGACTTG TAGCAGCTGA 120  
CCGAAGGTCC TGGAGATAC TGCTCTTTGG TGTAAATAAC TTGATATGTA CTGCGCTTCCT 180  
GCTTATGTGG TGCAGTTCTA CTAATAGTAT AGCTTTAACT GCCTATACCT ACCTGACCAT 240  
20 TTTTGATCTT TTTAGTTTAA TGACATGTTT AATAAGTTAC TGGGTAACAT TGAGGAAACC 300  
TAGCCCTGTC TATTCATTTG GGTTTGAAAG ATTAGAAGTC CTGGCTGTAT TTGCTCCAC 360  
AGTCTTGGCA CAGTTGGGAG CTCTCTTTAT ATTAAAGAA AGTGCAAGAC GCTTTTGGGA 420  
ACAGCCCGAG ATACACACGG GAAGATTATT AGTTGGTACT TTTGTGGCTC TTTGTTTCAA 480  
CCTGTTCAGC ATGCTTTCTA TCCGGAATAA ACCTTTTGCT TATGTCTCAG AAGCTGCTAG 540  
TAGCAGCTCG CTCAAGAGC ATGTTGCAGA TCTTAGTCGA AGCTTGTGTG GAATTATTCC 600  
25 GGGACTTAGC AGTATCTTCC TTCCCGGAAT GAATCCATT GTTTTGATTG ATCTTGCTGG 660  
AGCATTTGCT CTTTGTATTA CATATATGCT CATTGAAATT AATAATTATT TTGCGGTAGA 720  
CACTGCCTCT GCTATAGCTA TTGCCTTGAT GACATTTGGC ACTATGTATC CCATGAGTGT 780  
GTACAGTGGG AAAGTCTTAC TCCAGACAAC ACCACCCCAT GTTATTGGTC AGTTGGACAA 840  
ATCATCAGA GAGTATCTA CTTAGATGG AGTTTATAGAA GTCGAAATG AACATTTTTC 900  
30 GACCCTAGT TTTGGCTCAT TGGCTGGATC AGTGCAATGA AGAATTCGAC GAGATGCCAA 960  
TGAACAAATG GTTCTTGCTC ATGTGACCAA CAGGCTGTAC ACTTAGTGT CTACTCTAAC 1020  
TGTTCAAATT TTCAGGATG ACTGCTTAG GCCTGCCTTA TTGCTGGGCG CTGTTGCAGC 1080  
CAATGCTCTA AACCTTTTCAG ATCATCACT AATCCCAATG CCTCTTTTAA AGGGTACTGA 1140  
TGATTTGAAC AAGTTTACAT CAATCCAGC TAAACCTAGT AGTCCACCTC CAGAATTTTC 1200  
35 ATTTAACACT CCGGGGAAA AATGTGACCC AGTTATTCTT CTAACACAC AAACAGGCC 1260  
TTATGTTTCT GGTCTCAATC ATGGACACAC ACCCTACAGC AGCATGCTTA ATCAAGGACT 1320  
TGGAGTTCCA GGAATTTGAG CAATCAAGG ATTGAGGACT GGTTTTACAA ATATACCAAG 1380  
TAGATATGGA ACTAATAATA GAATTGGACA ACCAAGACCA TGATAGACTC TAACTTATT 1440  
40 TTATAAGGAA TATTGACTCC TTGGCTTCCA ATTTATTAG TAATCCAAC TTGCATTGAC 1500  
TGTTTAATCA TTTACTCTAA ATGTTAGATA ATAGTAGTCT TGTTACATT TCATGAAACC 1560  
TATGAAACTA TTTTCTTAA AATGTATTG GTGACAGTGA AATCCTCGTA AATGTTAAAG 1620  
GCTTTAAATA GGCTCTCTT AGAAATATGT TTTCTTAA AATTGATTGT GGTATCTTTG 1680  
GTTTTGTAGT TGACTGCACT GTGATGTGAC CTTACCTTTA TAAGAGCCAC TTGATGGAGT 1740  
45 AGATCTGTCA CATTACTAAG ATACGATATT TCTTTTCTT TCCGAGACGG AGTCTTGCTC 1800  
TGCCACTGTG CCGCGCAAT ACATTATTAT TAACCTAAGG CTGTACTTTA TTAAGGCTTC 1860  
CTTAGTTTTT GTTTTGTGTT GTTTTGTGAG ATGGAGTCTC ACTCTGTGCG CCAGGCTGGA 1920  
ATGCACTGGC ATGATCTCAG CTCAGTCCAA CCTCTGCTC CTGAGTTCAA ATGATTCTCC 1980  
TGCTCTAGCC TCCGAGTAG CTGGGATTAC AGGCACCTGC CACCAGCCCG AGCTAATTTT 2040  
50 TGATATTTTA GTAAAGACGG GGGATTTCAC CATGTTGGCC AGGCTGTGCT TGAACCTCTG 2100  
ACCTCATGAT CCACCCACCT TAGCCTCCCA AAGTGCTGGG ATTAGGTGTG AGCCACCGCA 2160  
CCTGGCCGAT ATTTCTTTA ATGAAATTTA TAAATATGCT TCTTGAATAA TACACATTTT 2220  
GGGAAAGGGA AAAATGCTG TTCAAAAAGT AAAGGTCTCT TTTATAGCTT TTCCAACTT 2280  
AATTGCTAAA TTTTCTTTG AGGTTCTCTT GAATTATGTC TTACAACTA AAAGCAAAAA 2340  
55 TTTTATGACG AAATTTTGA ATACATTCTA TCTAGCACAA TTTGAATTTT TAATTATCAA 2400  
GATTTTGTG AAAGTTTCTC TCCTTAAAA ATTTTAGTAC ATTTGTAAT

Seq ID NO: 483 Protein sequence  
Protein Accession #: BAB70980.1

1 11 21 31 41 51  
60 MGTIHLFRKP QRSFFGKLLR EFRIVAADRR SWKILLFVGI NLICTGFLLM WCSSTNSIAL 60  
TAYTYLTIFD LFLMTCLIS YWVLRKPSV VYSFGFERLE VLAVPASTVL AQLGALFILK 120  
65 ESAERFLEQP EIHTGRLLVG TFVALCFNLF TMLSIRNKPF AYVSEAASTS WLQEHVADLS 180  
RSLCGIIPGL SSIFLPRMNP FVLIDLGAFA ALCITYMLIE INNYFAVDTA SALAIALMTF 240  
GTMYPMSVVS GKVLQTTTPP HVIGQLDKLI REVSTLDGVL EVRNEHFNTL GFGSLAGSVH 300  
VRIIRDANBQ MVLAVHTNRL YTLVSTLTQV IPKDDWIRPA LLSGPVAANV LNFSDHHRVIP 360  
MPLLKGTDDL NPVTSTPAKP SSPPEFSFN TPGKNVNPVI LLNTQTREPYG FGLNHGHTPY 420  
70 SSMLNQGGLV FGIGATQGLR TGFTNIPSRV GTNNRIGQPR P

Seq ID NO: 484 DNA sequence  
Nucleic Acid Accession #: FGENESH predicted  
Coding sequence: 1..900

1 11 21 31 41 51  
75 ATGCGCGCGC GGGAGCTGAG CGAGGCGGAG CCGCCCGCGC TCCGGGCCCC GACCCCTCCC 60  
CCGCGCGCGC GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120  
80 CCGCTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180  
CGGCCCACTG CGCTGACAC CTTCTCTGGT ACGTACGTTT AATCGCCCGT CGGCGCGCT 240  
GGCTCGGGCG GGGCTGTGCA CCGGGGAGCT GGGCGGGGCG TCTCGGCGGG AGGGCGCAGA 300  
GGACCCCGGG GAGGAGACTG GAGCAGGCCG CGAGGTGGCG CTGGTGGCG CAGGACGCT 360  
CTTCCTAAGT CAGGCTCTCC CCGCCCGCGC CCGTCAAGTC AAGTCTGCTT GGATGGAGCT 420  
85 CCGGTGCGCA TTGAGCTCTG GGCACAGCG GGCAGGAGG ATTTTGACCG ACTTGTGTTCC 480  
CTTTGCTACC CGGATACCGA TGTCTTCTG GCGTGTCTCA GCGTGGTGCA GCCAGCTCC 540  
TTTCAAAACA TCACAGAGAA ATGGCTGCCG GAGATCCGCA CGCAACACCC CCAGGCGCCT 600  
GTGCTGCTGG TGGCACCCA GGCAGCTG AGGGACGATG TCAACGTACT AATTGAGTCT 660

GACCAGGGGG GCGGGAGGG CCGCGTGCCC CAACCCAGG CTCAGGGTCT GGCCGAGAAG 720  
 ATCCGAGCCT GCTGCTACCT TGAGTGCTCA GCCTTGACGC AGAAGAACTT GAAGGAAAGTA 780  
 TTGACTCGG CTATTCTCAG TGCCATTGAG CACAAAGCCC GGCTGGAGAA GAAACTGAAT 840  
 GCCAAGGTG TGCGACCCCT CTCCGCTGCG CGCTGGAAGA AGTTCTTCTG CTTCGTTTGA

Seq ID NO: 485 Protein sequence  
 Protein Accession #: PGENESH predicted

1 11 21 31 41 51  
 | | | | | |  
 MPPRELSEAE PPFLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60  
 RPTALDTPSG TVVQSPVVRPR GCGGAVERGA GAGVSAGRRR GPRGGDWSRP RGGAGAAQDA 120  
 LPNSGSPRPA PAVQVLVDGA PVRIELWDTA QGEDFDRLRS LCYPDTDVFL ACFSVVPSS 180  
 FQNTTEKWLFP EIRTNPQAP VLLVGTQADL RDDVNVLIQL DQGGREGPVP PQPAQGLAEK 240  
 IRACCYLECS ALTQKNLEKEV FDSAILSIAIE HKARLEKLN AKGVRTL SRC RWKFPFCFV

Seq ID NO: 486 DNA sequence  
 Nucleic Acid Accession #: XM\_063832.2  
 Coding sequence: 1..711

1 11 21 31 41 51  
 | | | | | |  
 ATGCCGCGCG GGAAGCTGAG CGAGGCGGAG CCGCCCCCGC TCCGGGCCCC GACCCCTCCC 60  
 CGCGCGCGCG GTAGCGCGCC CCCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120  
 CGCGTGCGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180  
 CGGCCCACTG CGCTGGACAC CTCTCTGTG CAAGTCTGGT TGGATGGAGC TCGGTGCGC 240  
 ATTGAGCTCT GGGACACAGG GGGACAGGAG GATTTTGACC GACTTCGTTC CCTTTGCTAC 300  
 COGATACCG ATGTCTTCTT GCGTGCTTTC AGCGTGGTGC AGCCACGCTC CTTTCAAAAC 360  
 ATCAGAGAGA AATGCGTGCC CGAGATCCGC AGGCACAACC CCCAGGCGCC TGTGCTGCTG 420  
 GTGGGCACCC AGGCGGACCT GAGGGACGAT GTCAACGTAC TAATTTCAGCT GGACGAGGGG 480  
 GCGCGGGAGG CGCCGCTGCC CCAACCCAG GCTCAGGGTC TGGCCGAGAA GATCCGAGCC 540  
 TGCTGCTACC TTGAGTGCTC AGCCTTGACG CAGAAGAACT TGAAGGAAGT ATTGACTCG 600  
 GCTATTCTCA GTGCCATTGA GCACAAAGCC CGCTGGAGA AGAACTGAA TGCCAAAGGT 660  
 GTGCGCACCC TCTCCCGCTG CCGCTGGAAG AAGTCTTCTT GCTTGTGTTG A

Seq ID NO: 487 Protein sequence  
 Protein Accession #: XP\_063832.1

1 11 21 31 41 51  
 | | | | | |  
 MPPRELSEAE PPFLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGYPARY 60  
 RPTALDTPSV QVLVDGAPVR IELWDTAGQE DFDRLRLSCY PDTDVFLACF SVVQPSFQFN 120  
 ITEKWLPEIR TKNPQAPVLL VGTQADLRDD VNVLIQLDQG GREGPVPQPO AQGLAEKIRA 180  
 CCYLECSALT QKNLEKEVDS AILSAIEHKA RLEKLNARG VRTL SRCRWK KFFCFV

Seq ID NO: 488 DNA sequence  
 Nucleic Acid Accession #: NM\_014398.1  
 Coding sequence: 64..1314

1 11 21 31 41 51  
 | | | | | |  
 GGCACCGATT CGGGGCTGCG CGGACTTCG CGCACGCTG CAGAACCTCG CCCAGCGCCC 60  
 ACCATGCCCC GCGAGCTCAG CGCGGCGGCC GCGCTCTCG CGTCCCTGGC CGTAATTTTG 120  
 CACGATGCGA GTCAAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT 180  
 ACTGCAGCAG CAACAGTACA GGACATAAAA AAACCTGTCC AGCAACGAGC TAAGCAAGCA 240  
 CCTCACCAAA CTTTAGCAGC AAGATTATG GATGGTCATA TCACCTTTCA AACAGCGGCC 300  
 ACAGTAAAAA TTCCAACAAC TACCCAGCA ACTACAAAA ACACCTGCAAC CACCAGCCCA 360  
 ATTACTTACA CCTCGGTGAC AACCCAGGCC ACACCCAACA ACTCACACAC AGCTCCTCCA 420  
 GTTACTGAAG TTACAGTCGG CCTAGCTTA GCCCTTATT CACTGCCACC CACCATCAAC 480  
 CCACCACTC ATACAGCTGG AACCAATTCA TCAACGTC CAACACCAAC TGGGAACACC 540  
 ACTCAACCCA GTAACAGAC CACCCCTTCCA GCAACTTTAT CGATAGCACT GCACAAAGC 600  
 ACAACCGGTC AGAAGCTTGA TCAACCCACC CATGCCCGAG GAACAACGGC AGCTGCCACC 660  
 AATACCAACC GCACAGCTGC ACCTGCCTCC ACGGTTCCTG GGCCCAACCT TGCACCTCAG 720  
 CCATGCTCAG TCAAGACTGG AATTTATCAG GTTCTAAAGC GAAGCAGACT CTGTATAAAA 780  
 GCAGAGATGG GGATACAGCT GATTGTTCAA GACAAGGAGT CGGTTTTTTC ACCTCGGAGA 840  
 TACTTCAACA TCGACCCCAA CGCAACGCAA GCCTCTGGGA ACTGTGGCAC CCGAAATCC 900  
 AACCTTCTGT TGAATTTTCA GGGCGGATTT GTGAATCTCA CATTTACCAA GGATGAAGAA 960  
 TCATATTATA TCAAGTGAAG GGGAGCCTAT TTGACCGTCT CAGATCCAGA GACAGTTTAC 1020  
 CAAGGAATCA AACATGCGGT GGTGATGTTT CAGACAGCAG TCGGGCATTC CTTCAAGTGC 1080  
 GTGAGTGAAC AGAGCCTCCA GTTGTGAGCC CACCTGCAGG TGAAAAACAAC CGATGTCCAA 1140  
 CTTCAAGCCT TTGATTTTGA AGATGACCAC TTTGGAAATG TGGATGAGTG CTCGCTGAC 1200  
 TACACAATTG TGCTTCTCTG GATTGGGGCC ATCGTGGTTC GTCTCTGCCT TATGGGTATG 1260  
 GGTGTCTATA AAATCGCCTT AAGTGTTCAA TCATCTGGAT ACCAGAGAAT CTAATTGTTG 1320  
 CCGCGGGGGA ATGAAAAATA TGGAAATTAG AGAACTCTTT CATCCCTTCC AGGATGGATG 1380  
 TTGGGAAATT CCCTCAGAGT GTGGGTCTTT CAAACAATGT AAACCAACCAT CTTCTATTCA 1440  
 AATGAAGTGA GTCATGTGTG ATTAAAGTTC AGGCAGCACA TCAATTTCTA AATACTTTT 1500  
 GTTTATTTTA TGAAAGATAT AGTGAGCTGT TTATTTTCTA GTTCTCTTTA GAATATTTTA 1560  
 GCCACTCAAA GTCAACATTT GAGATATGTT GAATTAACAT AATATATGTA AAGTAGAATA 1620  
 AGCCTTCAAA TTATAAACCA AGGGTCAATT GTAACATAAT CTACTGTGTG TGCATTGAAG 1680  
 ATTTTATTTT GATCTTGATC TTAACAAAGC CTTTGTCTTG TTATCAAAAT GACTTTCACT 1740  
 GCTTTTACIA TCTGTGTTT ATGGTTTCAT GTAACATACA TATTCCTGGT GTAGCACTTA 1800  
 ACTCCTTTTC CACTTTAAAT TTGTTTTTGT TTTTGGAGAC GGAGTTTCAC TCTTGTCAAC 1860  
 CAGGCTGGAG TACAGTGGCA CGATCTCGGC TTATGGCAAC CTCGCGCTCC CCGGTTCAAG 1920  
 TGATTCTCCT GCTTCAGCTT CCGAGTAGC TGGGATTACA GGCACACACT ACCAGCGCTG 1980  
 GCTAATTTT GTATTTTTAT TATAGACGGG TTTCAACATG TTGGCCAGAT TGGTCTTGAA 2040  
 CTCTTGACCT CAGGTGATCC ACCCACTCA GCCTCCCAA GTGCTGGGAT TACAGGCATG 2100  
 AGCCATTGCG CCGCGCCTTA AATGTTTTTT TTAATCATCA AAAAGAACAA CATATCTCAG 2160

GTTGTCTAAG TGTITTTATG TAAACCAAC AAAAAGAACA AATCAGCTTA TATTTTITAT 2220  
 CTGTGATGACT CCTGCTCCAG AATGCTAGA CTAAGAATA GGTGGCTACA GATGGTAGAA 2280  
 CTAAACAATA AGCAAGAGAC AATAAATATG GCCCTTAATT ATTAACAAGG TGCCAGAGTC 2340  
 TAGGCTAAGC ACTTTATCTA TATCTCATTT CATCTCACA ACTTAAGT GAATGAGTAA 2400  
 ACTGAGACTT AAGGGAAGTG AATCACTTAA ATGTCACCTG GCTAACTGAT GGCAGAGCCA 2460  
 GAGCTTGAAT TCACTTTGGT CTGACATCAA GGTCTTTGGT CTCTCCCTA CACCAAGTTA 2520  
 CCTACAAGAA CAATGACACC ACACCTGACC TGAAGGCTCA CACCTCATAC CAGCATACGC 2580  
 TCACCTTACA GGGAAATGGG TTTATCCAGG ATCATGAGAC ATTAGGGTAG ATGAAGAGAG 2640  
 AGCTTTGACG ATAACAAAT AGCCTATCCT TAATAAATCC TCCACTCTCT GGAAGGAGAC 2700  
 TGAGGGGCTT TGTAAACAT TAGTCAGTTG CTCATTTTAA TGGGATTGCT TAGCTGGGCT 2760  
 GTAAAGATGA AGGCATCAAA TAAACTCAAA GTATTTTAA ATTTTITGA TAATAGAGAA 2820  
 ACTTCGCTAA CCAACTGTTT TTTCTTGAGT GTATAGCCCC ATCTTGTTGT AACTTGCTGC 2880  
 TTTCTGACTT CATATCCATA TTTCTATTG TTTCTTTAT TCTGTAGAGC AGCCTGCCAA 2940  
 GAATTTTAT TCTGCTGTTT TTTTGTGTC TAAAGAAAGG AACTAAGTCA GGATGTTAAC 3000  
 AGAAAGTCC ACATAACCTT AGAATCTTA GTCAAGGAAT AATTCAAGTC AGCCTAGAGA 3060  
 CCAATGTTGAC TTTCTCATG TGTTCCTTA TGACTCAGTA AGTTGGCAAG GTCTGACTT 3120  
 TAGTCTTAAT AAAACATTGA ATTGTAGTAA AGGTTTTTGC AATAAAACT TACTTTTG

Seq ID NO: 489 Protein sequence  
 Protein Accession #: NP\_055213.1

1 11 21 31 41 51  
 MPRQLSAAAA LFASLAVILH DGSQMRKAP PETRDYSQPT AAATVQDIKK PVQQPAKQAP 60  
 25 HQTLAARFMD GHITFQTAAT VKIPTTTPAT TKNTATTSPI TYTLVTQAT PNNSTAPPV 120  
 TEVTVGPSLA PYSLPPTIIP PAHTAGTSSS TVSHTTGNT QPSNQTLPA TSLIALHKST 180  
 TQKQKQPTH APGTAAAHN TTRTAAPAST VPGTLPAPQ SSVKTIQVQ LNSRLCIAKA 240  
 EMGIQLIVD KESVFSPPRY ENIDPNATQA SGNCGTRKSN LLNFGGGFV NLTFTKDEES 300  
 30 YYISEVGAYL TVSDPETVYQ GIKHAVVMPQ TAVGHSFKCV SEQSLQLSAH LQVKTIDVQL 360  
 QAFDFEDDFH GNVDECSSDY TIVLPVIGAI VVGLCLMGMG VYKIRLRQS SGYQRI

Seq ID NO: 490 DNA sequence  
 Nucleic Acid Accession #: NM\_005409.3  
 Coding sequence: 94..378

1 11 21 31 41 51  
 TTCTTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAGCT GAAGTAGCAG 60  
 40 CAACAGCACC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGGCAT GGCTATAGCC 120  
 TTGGCTGTGA TATTGTGTGC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGACGC 180  
 TGCTTTTGCA TAGGCCCTGG GGTAAAAGCA GTGAAAGTGG CAGATATTGA GAAAGCCTCC 240  
 ATAATGTACC CAAGTAACAA CTGTGACAAA ATAGAAGTGA TTATTACCTT GAAAGAAAT 300  
 AAAGGCAAC AGTGCTTAAA TCCCAATCG AAGCAAGCAA GGCTTATAAT CAAAAAGTT 360  
 45 GAAAGAAAGA ATTTTAAAA ATATCAAAAC ATATGAAGTC CTGGAAGAGG GCATCTGAAA 420  
 AACCTAGAAC AAGTTTAACT GTGACTACTG AATGACAAG AATTCTACAG TAGGAACTG 480  
 AGACTTTTCT ATGGTTTGTG GACTTTCAAC TTTTGTACAG TTATGTGAAG GATGAAAGGT 540  
 GGGTGAAGG ACCAAAAACA GAAATACAGT CTTCCTGAAT GAATGACAAT CAGAAITCCA 600  
 CTGCCCAAG GAGTCCAGCA ATTAATGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT 660  
 50 GGTATACCATC GAGTATTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCAT 720  
 GCATTTCTAG CTGAGAGAAC CTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT 780  
 GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT 840  
 CTGTGCTTAC AGTGGAGACA TTGACATTAT TACTGGAGTC AAGCCCTTAT AAGTCAAAAG 900  
 CATCTATGTG TCGTAAAGCA TTCTCAAAAC ATTTTTCAT GCAATACAC ACTCTTTTCC 960  
 55 CCAATATCA TGATGACAT CAATATGTAG GGAACATTC TTATGCATCA TTTGGTTTGT 1020  
 TTTATAACCA ATCTATTAAA TGTAAATCAT AAAATGTACT ATGAAAAAAA TTATACGCTA 1080  
 TGGGATACCT GCAACAGTGC ACATATTTC TAACCAAAAT AGCAGCACCG GTCTTAATTT 1140  
 GATGTTTTT ACCTTTTATT CATTGAGATG TTTTGAAGCA ATTAGSATAT GTGTGTTTAC 1200  
 TGTACTTTTT GTTTTGATCC GTTTGTATAA ATGATAGCAA TATCTTGGAC ACATTTGAAA 1260  
 60 TACAAAATGT TTTTGTCTAC CAAAGAAAAA TGTTGAAAA TAAGCAAAAT TATACCTAGC 1320  
 AATCACTTTT ACTTTTGTGA ATTCTGTCTC TTAGAAAAAT ACATAATCTA ATCAATTTCT 1380  
 TTGTTCTATG CTATATACTG TAAATTTAG GTATACTCAA GACTAGTTTA AAGAAATCAA 1440  
 TGCAATTTTT TCTTAATAA ACTACCACAA CCTTTCTTTT TTAACAAAA AAA

Seq ID NO: 491 Protein sequence  
 Protein Accession #: NP\_005400.1

1 11 21 31 41 51  
 MSVKMAIAL AVILCATVVQ GPFMPKRGRC LCIGPGVKAV KVADISKASI MYPSNNCDKI 60  
 70 EVIITLKENK GQRCLNPKSK QARLIKKVE RKNP

Seq ID NO: 492 DNA sequence  
 Nucleic Acid Accession #: NM\_000577.1  
 Coding sequence: 41..520

1 11 21 31 41 51  
 GGCAAGAGGG GAAGACCTCC TGTCTATCA GGCCCTCCCC ATGGCTTTAG AGACGATCTG 60  
 80 CCGACCTCT GGGAGAAAAT CCAGCAAGAT GCAAGCCTTC AGAATCTGGG ATGTTAACCA 120  
 GAAGACCTTC TATCTGAGGA ACAACCAACT AGTTGCCGGA TACTTGCAAG GACCAATGT 180  
 CAATTTAGAA GAAAGATAC ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT 240  
 CCATGGAGGG AAGATGTGCC TGTCTGTGT CAAGTCTGGT GATGAGACCA GACTCCAGCT 300  
 GGAGGCACTT AACATCACTG ACCTGAGCGA GAACAGAAAG CAGGACAAGC GCTTCGCTT 360  
 85 CATCGCTCA GACAGTGGCC CCACCACAG TTTTGAAGTC GCCGCTGCCC CCGTTGGTT 420  
 CCTCTGCACA GCGATGGAAG CTGACCAGCC CGTCAGCCTC ACCAATATGC CTGACGAAGG 480  
 CGTCATGGTC ACCAAATTCT ACTTCCAGGA GGACGAGTAG TACTGCCCCG GCCTGCTCTG 540  
 TCCCATTCTT GCATGGCAAG GACTGCAGGG ACTGCCAGTG CCCCTGCCCC AGGACTCCCG 600

GCTATGGGG CACTGAGGAC CAGCCATTGA GGGGTGGACC CTCAGAAGGC GTCAACAACAA 660  
 CCTGGTCACA GGACTCTGCC TCCTCTTCAA CTGACCAGCC TCCATGCTGC CTCAGAATG 720  
 GTCTTTCTAA TGTGTGAATC AGAGCACAGC AGCCCCGTGA CAAAGCCCTT CCATGTGCGC 780  
 TCTGCATTCA GGATCAAAAC CCGACCACCT GCCCACTCTG CTCTCTCTT GCCACTGGCT 840  
 CTCTCTCTCT CATTCACACT TCCCATGCCC TGGATCCATC AGGCCACTTG ATGACCCCA 900  
 ACCAAGTGGC TCCCAACCTC TGTTTTACAA AAAAGAAAAG ACCAGTCCAT GAGGGAGGTT 960  
 TTTAAGGGTT TGTGAAAAT GAAAATTAGG ATTTTCATGAT TTTTTTTTTT CAGTCCCGCT 1020  
 GAAGGAGAGC CCTTCATTGG GAGATTATGT TCTTTOGGGG AGAGGCTGAG GACTTAAAAA 1080  
 ATTCCTGCAT TTGTGAAATG ATGGTGAAAG TAAGTGGTAG CTTTTCCTT CTTTTCCTT 1140  
 TTTTCTGTG ATGTCCCAAC TTGTAATAAT TAAAGTTAT GGTACTATGT TAGCCCCATA 1200  
 ATTTTTTTTT TCCTTTTAAA ACACCTCCAT AATCTGGACT CCTCTGTCCA GGCAGTCTGT 1260  
 CCCAGCTCC AAGCTCCATC TCCACTCCAG ATTTTITACA GCTGCTGCA GTACTTTACC 1320  
 TCCTATCAGA AGTTTCTCAG CTCCCAAGGC TCTGAGCAAA TGTGGCTCCT GGGGTTCTT 1380  
 TCTTCTCTG CTGAAGGAAT AAATTGCTCC TTGACATTGT AGAGCTTCTG GCACITGGAG 1440  
 ACTGTATGAA AAGATGGCTG TGCCTCTGCC TGTCTCCCC ACCAGGCTGG GAGCTCTGCA 1500  
 GAGCAGGAAA CATGACTCGT ATATGTCTCA GGTCCCTGCA GGTCCCAAGCA CCTAGCCTCG 1560  
 CTCTTGGCAG GTACTCAGCG AATGAATGCT GTATATGTTG GGTGCAAGT TCCTACTTTC 1620  
 CTGTGACTTC AGCTCTGTTT TACAATAAAA TCTTGAAAAA GCCTAAAAAA AAAAAAAA 1680  
 AAAAAAAA AAAAAAAA AAAAAAAA AAAAAA

Seq ID NO: 493 Protein sequence  
 Protein Accession #: NP\_000568.1

1 11 21 31 41 51  
 MALETICRPS GRKSSKMQAF RIWDVNQKTF YLRNNQLVAG YLQGPVNLE EKIDVVPFEP 60  
 HALEFLGHHG KMCLSCVKSQ DETRLQLEAV NITDLSENK QDKRFAPFIR DSGPTTSFES 120  
 AACPGWFLCT AMEADQFVSL TNMPDEGMV TKPYFQDE

Seq ID NO: 494 DNA sequence  
 Nucleic Acid Accession #: NM\_002081.1  
 Coding sequence: 222..1898

1 11 21 31 41 51  
 GGCTGCCGA GCGAGGTTT GGACCTGCA CCCCAGCGC CCGCGCGGC CGCCGCGCC 60  
 GGCTTTTGT GTCTCGCCT CCTCGCGCG CGCCGCTCT GGACCGCGAG CCGCGCGCG 120  
 CGGACCTTG GTCTCGCCT TCGCGGCGG GAACCTGCGA GGACCGCGC AGGATCCGAG 180  
 AGAGGCGCG GCGGGTGGC GGGGGCGCG CCGCGCGCG CATGGAGCTC CGGCGCGAG 240  
 GCTGTGGCT GCTATGTGCG GCGCGAGCG TGGTGGCTG CGCCGCGGG GACCGCGCA 300  
 GCAAGAGCGT CGTCTCGCG GAGGTGCGC AGATCTACG AGCCAGGGC TTCAGCTGA 360  
 CGGACGTGCC CAGGCGGAG ATCTCGGGT AGCACCTGCG GATCTGTCCC CAGGGCTACA 420  
 CCTGCTGCA CAGGAGATC GAGGAGAAC TGGCCAAAC CAGCCATGCC GAGCTGGAGA 480  
 CCGCGCTCG GGACAGCAG CGGTCTCTG AGGCCATGCT TGCCACCCAG CTGCGCAGCT 540  
 TCGATGACCA CTTCAGCAG CTGCTGAAG ACTCGGAGG GAGCTGCGC GCCACCTTCC 600  
 CGCGCGCTT CGGAGAGCTG TACACGAGA ACGCGAGGC CTTCCGGGAC CTGTACTCAG 660  
 AGCTGCGCT GTACTACCG GGTGCCAAC TGCACTGGA GGAGAGCTG GCGAGTTCT 720  
 GGGCGCGCT GCTCGAGCG CTCTTCAAG AGCTGCAACC CAGCTGCTG CTGCGTATG 780  
 ACTACCTGGA CTGCTGGGC AAGCAGGCG AGCGCTGCG GCGCTTGGG GAGGCGCGA 840  
 GAGAGCTGCG CCTGCGGGC ACCCGTGCCT TCGTGGCTGC TCGCTCTTT GTGCGGGCC 900  
 TGGCGGTGCG CAGGAGCTG GTCCGAAAG TGGCTCAGT CCGCTTGGG CCGGAGTGT 960  
 CGAGAGCTG CATGAAGCTG GTCTACTGT CTCACTGCT GGGAGTCCCC GCGCGCAGG 1020  
 CCTGCGCTGA CTATTGCTG AATGTGCTCA AGGGCTGCT TGCCAAACAG GCCGACCTGG 1080  
 ACGCGAGTG GAGGAACCTC CTGACTCCA TGGTGTCTAT CACCGACAG TTCTGGGTA 1140  
 CATCGGCTGT GGAAGGTGCT ATCGGAGCG TGCAACGCT GCTGGCGAG GCCATCAAC 1200  
 CCTTCAGGA CAACAGGAG ACCTCAGCG CCAAGTCTAT CCGGGCTGC GGAACCCCA 1260  
 AGGTCAACCC CAGGCGCTC GGGCTGAGG AGAAGCGCG CCGGGCAGG CTGCGCGCG 1320  
 GGGAGAGGCC ACCTTCAAG ACCTGAGGA AGCTGTCTC TGAAGCCAAG GCCAGCTCC 1380  
 GCGACGTCCA GGAATTCTG ATCAGCTCC CAGGGAACCT GTGAGTGAAG AAGATGGCCC 1440  
 TSAGCACTGC CAGTGATGAC CGCTGCTGA ACGGATGGC CAGAGGCGG TACCTCCCG 1500  
 AGGTCTATGG TGACGGCTG GCCAACAGA TCAACAACC CGAGGTGGAG GTGACATCA 1560  
 CCAAGCGGA CATGACCATC CGGAGCAGA TCATGCGACT GAAGATCAT ACCAACCGG 1620  
 TGGCAGGCG CTACAAAGCG AACGAGTGG ACTTCCAGGA CGCAGTGAC GACGCGAGG 1680  
 GCTCGGGCAG CGGTGATGCG TGTCTGATG ACCTCTGCG CCGGAGGTC AGCAGGAAGA 1740  
 GCTCCAGCTC CCGGAGCGCC TTGACCCATG CCTCCAGG CCGTCTCAG CAGGAAGGAC 1800  
 AGAAGACCTC GGCTGCCAGC TGCCCGCAG CCGCGACCT CTCTCTGCT CTCTCTCT 1860  
 TCCTGCGCTT TACAGTAGCC AGGCGCGGT GGGGTAACT GCGCCAGGC CCCAGGGACA 1920  
 GAGGCCAAG ACTGACTTTG CCAAAAATAC AACACAGAG ATATTTAATT CACTCAGCC 1980  
 TGGAGAGGCC TGGGTGGGA CAGGAGGCG CGGCGCTCT GAGCAGGGC AGCGCAGAG 2040  
 GTCCAGCCC CAGGCTGCG CTGCGCTGCC TTTCTGCTT TTAATTTGT ATGAGTCTCT 2100  
 CAGGTCACT GGGAGCCAGT GTGCCCAAAA GCCATGTATT TCAGGACCT CAGGGGCAAC 2160  
 TCGGCTGCG TAGCCTCCC CCGAGCTCCC TGACCGCGC CAGAAGCAGC CCCTCGAGG 2220  
 CTACAGAGA GGGCTCAAG CAACCGCTG GAGCCACAG CGAGCTGTG CCTTCTCTC 2280  
 CGCTCTCTC CACTGGACT CCGAGCAGG CCGACCAAG AGCCTTGGC CACCCCGAG 2340  
 CCTCCAGAGA AGCCCGCAC GGGCTGTCT GGTGTCCGC ATCCAGGTC TGGCAGAGC 2400  
 TCTGAGATGA TGCATGATG CCTCCCTCA GCGAGGCTG CAGAGCGCG CCGACCTCC 2460  
 CTGCGCTCT GAGGGGCCC AGGTCTGCA GGTGAGGCC TGAGACAGCA CCACTGCTGA 2520  
 GAGTCTGAG GACTGTCTCT CACAGACCC TGCACTGAGG GCGCTTCAAT CGCAGATGA 2580  
 GGGGCCACTG ACCACCTGC GCTTCTGCT GAGGAGGGA AGCTGGGCC AAAGGCCAG 2640  
 GAGGCGAGC TGGGTCTGC CAATGTGGC TGCCCTGCG ACACAGGCT CACAGGCGAG 2700  
 GCTTGTCTG GGTTCAGGCG TGTGGAGA CCGGAGGCG TGAGGAGCAG CAGGACCCG 2760  
 CCTGCTCCA TCCTCACCCA GATCAGGAAC CAGGGCTCC CTGTTACGCG TGACACAGT 2820  
 CAGGCTCAG AGTGACCTCT GGTCTGACCC TGCTCACAGG GATGTGCTG GCTGTGAGA 2880  
 CCGCGACTG CACAGGGAA TGCTAGGTC CCTTCCGAG CAGGCGACT GCACTGAGG 2940  
 GCAAGGAGC CTGATAGT AAGGGCTTT CCAACATGC ATCCATTAC TGACACTTC 3000  
 TGTCTTGT CATGAGAGC TGTCTGCTC TCCAGATGG CTTGAGGCG CCGCAGGCG 3060  
 CACTTGGAC CCTGTGACC TCCTGTCACT CACTGAGGCC ATCAGGCGC TGCCCAAGC 3120

CTGGAGGGGC CCTCTTCCC TCCTGTGCC CAGCTGCCAG GTGGCCCTGG GGAGGGGTGG 3180  
 TGTGTGTGTG GGAAGGGGTC CTGCAGGGGG AGGAGGACTT GGAGGGTCTG GGGGCAGCTG 3240  
 TCCTGAACCG ACTGACCCCTG AGGAGGGCCG TTAGTGCTGC TTTGCTTTTC ATCACCGTCC 3300  
 CGCACAGTGG ACGGAGGTCC CCGGTTGCTG GTCAGGTCCC CATGGCTTGT TCTCTGGAAC 3360  
 CTGACTTTAG ATGTTTGTGG ATCAGGAGCC CCCAACACAG GCAAGTCCAC CCCATAATAA 3420  
 CCTGTCCAGT GCCAGGGTGG GCTGGGGACT CTGGCACAGT GATGCGGGG GCCAGGACAG 3480  
 CAGCACTCCC GCTGCACACA GACGGCCCTAG GGGTGGCGCT CAGACCCAC CCTACGCTCA 3540  
 TCTCTGGAAG GGGCAGCCCT GAGTGGTCAC TGGTCAGGCG AGTGGCCAAG CCTGTGTGTG 3600  
 CTTTCTCCA CAAGTCCCC CCACCGCTCA GTGTAGCGG GTGACGTGTG TTTCTTTGAG 3660  
 TCCTTGTATG AATAAAGGC TGGAAACCTA AA

Seq ID NO: 495 Protein sequence  
 Protein Accession #: NP\_002072.1

1 11 21 31 41 51  
 MELRARGMWL LCAAAALVAC ARGDPASKSR SCGEVRQIYG AKGFSLSDVP QAEISGEHLR 60  
 ICPGGTYTCT SEMRENLANR SHAELETALR DSSRLQAML ATQLRSFDDH FOHLNDSER 120  
 TLQATPPGAF GELYTONARA PRDLYSELRL YYRGANLHLE ETLAEFWARL LERLFQQLHP 180  
 QLLEPDDYLD CLGQAEALR PFGEAPREL RLRATAPFAA RSPVQGLGVA SDVVRKVAQV 240  
 PLGPECSRVA MKLVYCAHCL GVPGARPCPD YCRNVLKGCL ANQADLDAEW RNLLDSMVL I 300  
 TDKFWGTSGV ESVIGSVHTW LAEAINALQD NRDITLAKVI QGCGNPKVNP QGPGPEEKRR 360  
 RGKLAPRERP PSGLLEKLVS EAKAQLRDVQ DFWISLPGTL CSEKNALSTA SDDRWNMGMA 420  
 RGRVLPFVMG DGLANQINNP EBEVDITKPD MTIRQQIMQL KIMTNRLRSA YNGNDVDFQD 480  
 ASDDGSGSGS

Seq ID NO: 496 DNA sequence  
 Nucleic Acid Accession #: NM\_001650.2  
 Coding sequence: 40..1011

1 11 21 31 41 51  
 GGGGAGGCA ATGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCACAGCA 60  
 AGGCGTGTGG GTAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCTTTCAAA 120  
 GGGGTCTGGA CTCAGCTTTT CTGGAAGACA GTCACAGCGG AATTTCTGGC CATGCTTATT 180  
 TTTGTTCTCC TCAGCGTGGG ATCCACCATC AACTGGGGTG GAACAGAAAA GCCTTTACCG 240  
 GTGACATGG TTCTCATCTC CCTTTGCTTT GGACTCAGCA TTGCAACCAT GGTGCACTGC 300  
 TTTGGCCATA TCAGCGGTGG CCACATCAAC CCTGCAGTGA CTGTGGCCAT GGTGTGCACC 360  
 AGGAAGATCA GCATCGCCAA GTCTGTCTTC TACATCGCAG CCCAGTGCCT GGGGGCCATC 420  
 ATGGAGCAG GAATCCTCTA TCTGGTCACA CCTCCAGTG TGGTGGGAGG CTGGGAGTGC 480  
 ACCATGGTTC ATGGAATCT TACCGCTGGT CATGGTCTCC TGGTTGAGTT GATAATCACA 540  
 TTTCAATTGG TGTTTACTAT CTTTGCCAGC TGTGATTCGA AACGAGCTGA TGTCACTGGC 600  
 TCAATAGCTT TAGCAATTGG ATTTTCTGTT GCAATTGGAC ATTATTATTC AATCAATTAT 660  
 ACTGGTGCCA GCATGAATCC CGCCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG 720  
 GAAACCAATT GGATATATTG GGTGGGCCC ATCATAGGAG CTGTCCTGCG TGGTGGCCTT 780  
 TATGAGTATG TCTTCTCTCC AGATGTTGAA TTCAAACGTC GTTTTAAAGA AGCCTTCAGC 840  
 AAAGCTGCCC AGCAAAACAA AGGAAGCTAC ATGGAGGTGG AGGACAACAG GAGTCAGGTA 900  
 GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGCAATG TGATTGACST TGACCGGGGA 960  
 GAGGAGAAAG AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAGTATG ACTAGAAGAT 1020  
 CGCACTGAAA CGACACAAGA CTCCTTAGAA CTGTCTCTAG ATTTCCTTCC ACCCATTAAG 1080  
 GAAACAGATT TGTATATAAT TAGAAATGTG CAGGTTTGTG GTTTCATGTC ATATTACTCA 1140  
 GTCTAAACAA TAAATATTTC ATAATTACA AAGGAGGAAC GGAAGAAACC TATTGTGAAT 1200  
 TCCAAATCTA AAAAAGAAA TATTTTAAG ATGTTCTTAA GCAAAATATAT ACCTATTTTA 1260  
 TCTAGTTACC TTTTATTAAC AACCATTATT AACCGTGTGT CAAGATTGGG TTAAGTCTTG 1320  
 CCTGACAGAA CTCAAAGACA CGTCTATCAG CTTATTCCTT CTCTACTGGA ATATTGGTAT 1380  
 AGTCAATTCT TATTGAATA TTTATCTAT TAAACTGAGT TTAACATGG C

Seq ID NO: 497 Protein sequence  
 Protein Accession #: NP\_001641.1

1 11 21 31 41 51  
 MSDRPTARRW GKCGPLCTRE NIMVAFKGVW TQAFWKAVTA EFLAMLIPVL LSLGSTINWG 60  
 GTEKPLPVDV VLISLCFGLS IATMVQCFGH ISGGHINPAV TVAMVCTRKI SIAKSVFYIA 120  
 AQCLGAIIGA GILYLVTPPS VVGGLGVTMV HGNLTAGHGL LVELIITPQL VFTIPASCD S 180  
 KRTDVTGSIA LAIGFVSAIG HLPAINYTGA SMNPARSFPF AVIMGNWENH WIYVVGPIIG 240  
 AVLAGGLYEV VPCPDVEFKR RFKEAFSKAA QQTGGSYMEV EDNRSQVETD DLILKPGVVH 300  
 VIDVDRGEEK KGKQDSGEVL SSV

Seq ID NO: 498 DNA sequence  
 Nucleic Acid Accession #: AB020684.1  
 Coding sequence: 1..1744

1 11 21 31 41 51  
 CCCCTTGTG ATTAATACAT TAAAAAGATT CAATCTTTAC CCTGAGGTAA TTTTGGCCAG 60  
 TTGGTACCGG ATTTATACCA AAATAATGGA CTGTATTGGT ATTCAAACCA AGATATGTTG 120  
 GACGGTTACC AGAGGAGAAG GACTCAGTCC TATTGAAAGC TGTGAAGGAT TGGGAGATCC 180  
 TGCTGTCTTT TAGTGTGCTG TAATTTTAT TTTAAATGGA CTAATGATGG CATTATTCTT 240  
 CATATATGGC ACATATTTAA GTGGCAGCCG ATTAGGAGGC CTGGTTACAG TGTGTGCTT 300  
 CTTTTCAAT CATGGAGAGT GTACCGGTGT AATGTGGACA CCACCTCTCC GTGAAAGCTT 360  
 CTCATATCCA TTTCTTGTTC TTCAGATGTT GCTAGTGACT CATATTCTCA GGGCTACAAA 420  
 ACTTTATAGA GGAAGCTTGA TTGCACTCTG CATTTCCAAT GTATTTTCA TGCTTCTTGT 480  
 GCAGTTTGCT CAGTTTGTAT TTCTTACTCA GATTGCATCA TTATTTCAG TATATGTGT 540  
 CGGGTACATT GATATATGTA AATTACGAA GATCAITTTAT ATACACATGA TTTCTCTTGC 600  
 ACTTTGTTT GTTTTGATGT TTGGGAACCT AATGTTATTA ACTTCTTATT ATGCTTCTTC 660  
 TTTGGTAATT ATTTGGGGA TTTCTGCAAT GAAACCATAT TTTCTGAAAA TAAATGTATC 720

	TGAACTTAGT	TTATGGGTTA	TTCAAGGATG	TTTTTGGTTA	TTTGGAACTG	TCATACTTAA	780
	ATACTTGACA	TCTAAAATTT	TTGGTATTGC	AGATGACGCT	CATATTGGCA	ACTTACTAAC	840
	ATCAAAATTC	TTTAGTTATA	AGGATTITGA	TACTTTATTG	TATACCTGTG	CAGCGGAGTT	900
5	TGACTTTATG	GAAAAGAGAG	CTCCACTGAG	ATACACAAAG	ACATTATTGC	TTCCAGTTGT	960
	CTTGTAGTGT	TTTGTGTGTA	TTGTTAGAAA	GATTATTAGT	GATATGTGGG	GTGTCTTAGC	1020
	TAAACACAG	ACACATGTAA	GAAAACACCA	GTTTGATCAT	GGAGAGCTGG	TTTACCATGC	1080
	ATTGCAATTG	TTAGCATATA	CAGCCCTTGG	TATTTTAATT	ATGAGACTAA	AACTCTTCTT	1140
	GACACCACAC	ATGTGTGTTA	TGGCATCACT	GATCTGCTCA	AGACAGCTAT	TTGGATGGCT	1200
10	CTTTTGCAAA	GTACATCTCG	GTGCTATTGT	GTTTGCTATA	TTAGCAGCAA	TGTCAATACA	1260
	AGGTTTCAGCA	AATCTGCAAA	CCCAAGTGAA	TATTGTAGGG	GAGTTCAGCA	ATTTGCCCCA	1320
	AGAAGAAGCT	ATAGAATGGA	TCAAATATAG	TACTAAACCA	GATGCAGTGT	TTGCGGGTGC	1380
	CATGCCACAG	ATGGCAAGTG	TTAAGCTCTC	TGCACCTCGG	CCCATTTGTA	ATCATCCACA	1440
	TTATGAAGAC	GCAGGCTTGA	GAGCCAGAAC	AAAAATAGTA	TACTCAATGT	ATAGTCGGAA	1500
15	AGCAGCCGAA	GAACTGTAAG	GAGAACTGAT	AAAGTTAAAA	GTGAACATT	ACATTCTAGA	1560
	AGAGTCATGG	TGTGTAAAGAA	GATCCAAAGCC	TGGTTGCAGT	ATGCCCTGAA	TTTGGGATGT	1620
	AGAAGATCCT	GCCAATGCTG	GGAAAACCTCC	CTTATGTAAC	CTCTTGGTGA	AGGATTCCAA	1680
	ACCTCACITTC	ACCACGTGAT	TCCAGAACAG	TGTTTACAAA	GTCTTAGAAG	TTGTAAAAGA	1740
	ATGACTGCTA	CATGACCTGC	TGCCTAAGGA	GAACATACAT	TGTAAATGTT	TTAATGTTTT	1800
20	GCTAAGTCAT	GTGTGTTTCA	TATCCCAAAA	ACTTTTATAG	GTAACGTGTT	TCAAATAGAA	1860
	AACGTTTTAT	TTGGTCAATT	TGAATGTCAT	TCTAATTATA	AAAAATGACT	ACACCTTTAT	1920
	CAATTGGTTA	GTGTTTCAAT	GCACCTTTTA	AAATTGCTTA	TGCAAAATGAG	TATATGCTTG	1980
	TACTTGACCT	TAATATTGTT	GCTAAAGTGA	GCAAAGCTAC	CTGTATAAAG	AAAAACACAGT	2040
	GGGTTGTGAC	AAGGATGACA	TGAAAATACA	GGACAATTCT	GACAATGTAG	GGGCTGATTT	2100
25	TATAGTGTAA	GAACTATTAA	TGCCCTTTC	TTCTTTTTC	TGCCCTTTC	TCTTGTCTTT	2160
	TGGACATTTT	AGTGTATGTA	AGTTCCTCGG	TCAATGTCAG	CCCTGTCATC	AACITGAGTT	2220
	ACAGTAGATG	GGGACAGCAT	GGAGTGTGTT	CTATATAAAA	CTATCTGTTT	GTCTTACTTC	2280
	CTGTGCGCT	TTTTGTTCTC	TGTTCTCTTG	TTAATGAAGC	TTTCTCTGCC	CATTATTAAAT	2340
	CCAACTCTT	GGACCTTGTG	GTAGGAAAT	TCCCTTAACT	TCCAGCCATA	TGGCATTATC	2400
30	GTGTCTCTTT	CTCTCTCTCT	CTTGTCTCTC	CTCTCTCTCT	CTTCCCATTA	TTTTCTGTCA	2460
	AATAAGTACT	GTCTTACTCAT	TAGTGTGCTT	ATCAAGTACT	TATCTTGGT	TTTAAAAAAA	2520
	ATTAATGTGA	ACTGTATTCT	TCTCATTTTT	AGCATTTATC	AAATGTTTAT	ATTTTAAATAC	2580
	CTTTAAACCA	CTTTAAAGTT	TTTTCATGTT	TAATTATAGT	TTTAAAGAAA	ACTATTTTGA	2640
	ACRAACCCAA	ATATAGTGCA	TCTAGAAACT	AATGTATATT	TGATTAGACA	TCATTATTAG	2700
35	TGGAACAGTA	GACTGTAGTA	CATGGTAATT	TTTCTTTTAC	TATTAAGATA	CAATAAAACA	2760
	TGACTAAATT	TGCTGTCAAA	AATGTAAAGA	ATAATGATAA	ATGGAGTTTT	TTATATTTTA	2820
	CTTTAAAGAT	TGCTGTCTCT	TAATAAGACA	AAGCCTTAAG	CCTTATGTTA	TAATTTTGGT	2880
	CTTAAAAACC	ATCATTTTCAG	TATAAGGAAT	AAGTATATTT	CGTCTCCTC	TTTAGTTTTT	2940
	TTCTTCTTAT	TATTTTCTAT	TTTGAATAAT	TTCTACACCT	TCTTTGAATT	CCTTGTATGA	3000
40	ATTTTGTGTT	CTTAGAAGTT	AATTTGTGTG	AAATGAGATT	CTTCAAAACG	ATGAAACCTC	3060
	ATAGCTCTGA	GAAAAGGTTT	TAGGGTTTTA	AATTCIAAGC	AAAGCGTGAC	TATGGCTGAC	3120
	AGACTACACA	TTTAATTATA	CAGCTTCTCT	TTCTTAACCA	CAGGCAGATT	AACCTCATTT	3180
	TGGATTGTCC	TTACAGACCT	AGTCTCTCAG	CATGGTTTTT	GGTGGCCACT	CCTGGAAGCC	3240
45	GTGTGTTCCCT	TTCTACCTTC	TTACCAGAGC	CCAAGGGCAG	GCCTGGTCCC	GGGGAAGCAG	3300
	CAGCTTGCTG	ACATAAGTCA	GCTGCAAGG	CTGAGGAGTG	TGCCCTCAGA	GAAGCACCGC	3360
	CCCCAGTCT	TGTGCGACGG	CCTAGAGCCG	CAGCTCCAG	GGATGCTCCT	TCCCTGGAGG	3420
	CAGCCACGGA	GAGGAGCTCT	GGCAGCGTTC	TTCAGATTG	TGSCCACTGT	TTCTCATTTG	3480
	CTGGTTGACT	GTTTTTTATT	CTTAGGCTTT	TGCTAGTTTT	AGAAAATAGG	GAAGCAGCCC	3540
	TGGAATTGTG	GATTAAGAGC	AACATTGAG	CGATGATGCA	CAACAGTCCA	GGAAAATGGG	3600
50	OGGTGACAC	TTGAGGCTGA	GGATGGGAGT	TGACATGAGC	AGGGAGAGGG	AGGTGCGCGC	3660
	TGCTTATCTG	TGATTGTTGC	TCACCTGAGT	GTGGCTGATT	GTGTACATCC	AGCAGTTACA	3720
	ATTTTAAAAA	ATTATACTTT	TACATTATT	TTATATTTTT	CTCACCCCCA	GTAATTTCTT	3780
	TCCAAGAGAG	TTACATGTA	ATAAGTAGAA	ATTCTGTATA	GGAAAAAAGC	ATTAAAAATA	3840
	CTATTATAAC	TGCTTCAATT	GCTGGGAACC	ATTAAAGATA	ATATAAATTA	GCITTTTCCA	3900
55	GAAGGATCCT	TTGTAGCAG	TGTTATGAA	TGTAACCCCC	AGCAAAATAT	GGCTATATAT	3960
	TAGGGGAGCC	AGTTTGGAGC	AGAGGCCCTGA	AGTCCCTGC	TATGCAGCCG	TGGCCACAGC	4020
	TGCGAGCCCA	AGCACTGTGG	AGCATCCACA	CCTTTGATGG	CAATGCAGAT	TGGTAGCAGG	4080
	TTCCATAGGC	GTACAAAACA	GTATTAAAGC	TCAGTGTGTT	GCAATTTGTT	AGCATTTACA	4140
	AAATATTTTG	CTTTAGTATG	AGGAAAGTAA	GGATGGGCAA	AGAAGCGATC	AAAAATAGCTA	4200
60	TTGCTACAAC	ATTTTCGAAA	ACAAAGTTGG	GGCTGTATTT	CTTTAAAAAG	ATAAGCCTCT	4260
	AAAAATGCTT	GGCAAAAAAA	ATATAGTGTT	AAAAATAGCC	AGTGATATTA	ATGAGAAAAT	4320
	GAAAGTATGT	ATCAGGAATA	AAGTGATATT	GCAATAGGAGT	ATTGTATTTT	TATGAATTTT	4380
	ATGCCAGTTG	TTTACATGTA	CTATATATGT	TAAATTAAAA	AAAATCATGA	GAAATG	

Seq ID NO: 499 Protein sequence

Protein Accession #: BAA74900.1

	1	11	21	31	41	51	
70	PLVINTLKR	NLYPEVILAS	WYRIYTKIMD	LIGIQTICW	TVTRGEGLS	IESCEGLGDP	60
	ACFYVAVIFI	LNGLMMALFF	YGTLYSGSR	LGGLVTVLCF	PFNEGECTRV	MWTPPLRESF	120
	SYFPLVLQML	LVTHILRATK	LYRGSILALC	ISNVFMLEFW	QPAQFVLLTQ	IASLPAVVVV	180
	GYIDICKLRK	IYIHMISLA	LCPVLMFGNS	MLLTSSYASS	LVIIWIGILAM	KPHFLKINVS	240
	ELSLMWIGQC	PWLPFTVILK	YLTSKIPIGIA	DDAHIGNLLT	SKFPYSKDFD	TLIYTCAAEF	300
75	DFMEKETPLR	YTKTLLLPV	LVVFAIVRK	IISDMGVLA	KQTHVRKHQ	FDHGLVYHA	360
	LQLLAYTALG	ILIMRLKLF	TPHMCVMASL	ICSRQLFGWL	PCKVHPGAIV	FAILAAMSIIQ	420
	GSANLQTQWN	IVGEPSNLFP	EELIEWIKYS	TKPDVAFAGA	MPTMASVKLS	ALRPVNHPPH	480
	YEDAGLRART	KIVYSMSYRK	AAEEVKRELI	KLKVNYYLE	ESWCVRRESKP	GCSMPEIWDV	540
	EDPANAGKTP	LCNLLVKDSK	PHFTTVFGNS	VYKVLVVKE			

Seq ID NO: 500 DNA sequence

Nucleic Acid Accession #: NM\_001276.1

Coding sequence: 127..1278

	1	11	21	31	41	51	
85	AGTGGAGTGG	GACAGGTATA	TAAAGGAAGT	ACAGGGCCTG	GGGAAGAGGC	CCTGTCTAGG	60
	TAGCTGSCAC	CAGGAGCCGT	GGGCAAGGGA	AGAGGCCACA	CCCTGCCCTG	CTCTGCTGCA	120

GCCAGAATGG GTGTGAAGGC GTCTCAAACA GGCTTTGTGG TCCTGGTGCT GCTCCAGTGC 180  
 TGCTCTGCAT ACAAACTGGT CTGCTACTAC ACCAGCTGGT CCCAGTACCG GGAAGGCGAT 240  
 GGGAGCTGCT TCCAGATGCT CCTTGACCGC TTCTCTGTGA CCCACATCAT CTACAGCTTT 300  
 5 GCCAATATAA GCAACGATCA CATCGACACC TGGGAGTGA ATGATGTGAC GCTCTACGGC 360  
 ATGCTCAACA CACTCAAGAA CAGGAACCCC AACCTGAAGA CTCTCTGTCT TGTCCGAGGA 420  
 TGGAACTTTG GGTCTCAAAG ATTTTCCAAG ATAGCCTCCA ACACCCAGAG TCGCCGGAAT 480  
 TTCATCAAGT CAGTACCGCC ATTCTGCGC ACCCATGGCT TTGATGGGCT GGACCTTGCC 540  
 TGGCTCTACC CTGGACGGAG AGACAAACAG CATTTTACCA CCCTAATCAA GGAATGAAG 600  
 10 GCGAATTTA TAAAGGAAGC CCAGCCAGGG AAAAAGCAGC TCCTGCTCAG CGCAGCACTG 660  
 TCTGCGGGGA AGGTCAACAT TGACAGCAGC TATGACATTG CCAAGATATC CCAACACCTG 720  
 GATTTCATTA GCATCATGAC CTACGATTTT CATGGAGCCT GGCGTGGGAC CACAGGCCAT 780  
 CACAGTCCCC TGTTCGAGG TCAGGAGGAT GCAAGTCTGT ACAGATTGAG CAACACTGAC 840  
 TATGCTGTGG GGTACATGTT GAGGCTGGGG GCTCCTGCCA GTAAGCTGGT GATGGGCATC 900  
 15 CCCACCTCG GGAAGAGCTT CACTCTGGCT TCTTCTGAGA CTGGTGTGG AGCCCCAATC 960  
 TCAGAGCCGG GAATTCAGG CCGGTTCCAC AAGGAGGAGG GGACCCCTGC CTACTATGAG 1020  
 ATCTGTGATC TCCTCCGCGG AGCCACAGTC CATAGAACCC TCGGCCAGCA GGTCCCTAT 1080  
 GCCACCAAGG GCAACCAAGT GGTAGGATAC GACGACCAGG AAAGCCTCAA AAGCAAGGTG 1140  
 CAGTACTGTA AGGATAGGCA GCTGGCAGGC GCCATGGTAT GGGCCCTGGA CTTGATGAC 1200  
 20 TTCCAGGGCT CCTTCTGCGG CCAGGATCTG CGCTTCCCTC TCACCAATGC CATCAAGGAT 1260  
 GCACCTGCTG CAACGTAGCC CTCTGTCTG CACACAGCAG GGGGGCCAAG GATGCCCGCT 1320  
 CCCCTCTCG CTCCAGCTCG CCGGAGCCCT GATCACCTGC CCTGCTGAGT CCCAGGCTGA 1380  
 GCCTCAGTCT CCTCCCTTG GGGCCTATGC AGAGGTCCAC AACACACAGA TTTAGCTCA 1440  
 GCCCTGGTGG GCAGAGAGGT AGGGATGGGG CTGTGGGGAT AGTGAGGCAT CGCAATGTAA 1500  
 25 GACTCGGGAT TAGTACACAC TTGTGTATGA TTAATGAAA TGTTTACAGA TCCCCAAGCC 1560  
 TGGCAAGGGA ATTTCTTCAA CTCCTGCCCC CCTAGCCCTC TTATCAAAG GACACCATTT 1620  
 TGGCAAGCTC TATCAACCA GAGCCAAACA TCCTACAAGA CACAGTGACC ATACTAATTA 1680  
 TACCCCTGCG AAAGCCAGCT TGAACCTTC ACTTAGGAAC GTAATCGTGT CCCCTATCCT 1740  
 ACTTCCCTCT CCTAATTCAC CAGCTGCTCA ATAAAGTACA AGAGTTTAA AGTGTGTGG 1800  
 30 CGCTTTGCTT TGGTCTATCT TTGAGCGCCC ACTAGACCCA CTGGACTCAC CTCCCCCATC 1860  
 TCTTCTGGGT TCCTCTCTCT GAGCCTTGGG ACCCTTGAGC TTGCAGAGAT GAAGGCCGCC 1920  
 ATGTT

Seq ID NO: 501 Protein sequence  
 Protein Accession #: NP\_001267.1

1 11 21 31 41 51  
 MGVKASQTGF VVLVLQCCS AYKLVYYTS WSQYREGDGS CFPDALDRFL CTHIIYSPAN 60  
 40 ISNDHIDTWE WNDVTLYGML NTLKNRNPFL KTLISVGGWN FGSQRFPSKIA SNTQSRRTFI 120  
 KSVFPPLRTH GFDGLDLAWL YPGRRDRQEF TTILKEMKAE FIKEAPGPKK QLLLSAALSA 180  
 GKVTIDSSVD IAKISQHLDF ISIMTYDFHG AMRGTTHHHS PLFRGQEDAS PDRFSNTDYA 240  
 VGYMLRLGAP ASKLVMIPT FGRSFTLASS ETGVGAPISG PGIFGRFTKE AGTLAYYEIC 300  
 DFLRGATVHR TLGQVVPYAT KGNQWVGYYD QESVSKVQY LKDRQLAGAM VWALDLDDFQ 360  
 45 GSFGQDLRF PLTNAIKDAL AAT

Seq ID NO: 502 DNA sequence  
 Nucleic Acid Accession #: NM\_006474.1  
 Coding sequence: 181..669

1 11 21 31 41 51  
 GCTGCTAGG GTCTGGAAG CTGCGGCACC CTCCTCTCC GGGGCTCCTG CTCCACCCC 60  
 TCGGGCCCCC CCACCGTCCG GCTCCTCCAG GCTGGGCTGT TGGCGCGGT GCTTTTAATT 120  
 55 TTCCCCCAGC TCAGAATCTT GCTGCTCGGC CCCAGGAGA GCAACAACCT AACGGGAAGC 180  
 ATGTGAAGG TGTCAGCTCT GCTCTTCGTT TTGGGAAGCG CGTGCCTCTG GGTCTTGCA 240  
 GAAGGAGCCA GCACAGGCCA GCCAGAAGAT GACACTGAGA CTACAGGTTT GGAAGGCGGC 300  
 GTTGCCATGC CAGGTGCCGA AGATGATGTG GTGACTCCAG GAACCAAGCA AGACCGCTAT 360  
 AAGTCTGGCT TGCAACTCTT GGTGGCAACA AGTGTCAACA GTGTAACAGG CATTGCAATC 420  
 60 GAGGATCTGC CAATCTCAGA AAGCACATC CACGCGCAAG AACAAAGTCC AAGCGCCACA 480  
 GCCTCAAGC TGGCCACCAG TCACTCCAGS GAGAAAGTGG ATGGAGACAC ACAGACAACA 540  
 GTTGAGAAAG ATGTTTGTG AACAGTGACC CTGGTTGGAA TCATAGTTGG GGTCTTACTA 600  
 GCCATCGGTT TCATTGGTGG AATCATCGTT GTGGTTATGC GAAAAATGTC GGGAGGTAC 660  
 TCGCCCTAAA GAGCTGAAG GTTACGCCCT GCTTGCCAAC GTGCTTTAAA AAAAGACCGT 720  
 65 TCTGACTCT GTGGCCCTGT CCTGAGCTC GTGGGGAGAA GATGACCCCT GGAACATTGT 780  
 CGGGCCCAAT CAGATTCCAC GGTGACTTTC CGTTTGCCAA ATTAACGAG GAAAGACCTT 840  
 TCACCAGATT TGGTCTTAA ACTTT

Seq ID NO: 503 Protein sequence  
 Protein Accession #: NP\_006465.1

1 11 21 31 41 51  
 MWKVSALLFV LGSASLWVLA EGASTGQPED DTETTGLEGG VAMPGAEDDV VTPGTSEDRY 60  
 75 KSGLTLLVAT SVNSVTGIRI EDLPTSESTV HAQEQSFSAT ASNVATSHST EKVDGDTQTT 120  
 VEKDELSTVT LVGIIVGVLL AIGFIGGIIV VMRKMSGRY SP

Seq ID NO: 504 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 62..895

1 11 21 31 41 51  
 CACTGCTCTG AGAATTTTGT AGCAGCCCCC AACAGGCTGT TACTTCACTA CAACTGACGA 60  
 85 TATGATCATC TTAATTTACT TATTTCTCTT GCTATGGGAA GACACTCAAG GATGGGGATT 120  
 CAAGGATGGA ATTTTTCATA ACTCCATATG GCTTGAACGA GCAGCCGGTG TGTACACAG 180  
 AGAAGCACGG TCTGGCAAT ACAAGCTCAC CTACGCAGAA GCTAAGGCGG TGTGTGAATT 240  
 TGAAGCGGCG CATCTGSCAA CTTACAAGCA GCTAGAGGCA GCCAGAAAAA TTGGATTTC 300



5  
10  
15  
20

TGTCGTGCT GCTGGATGGA TGGCTAAGGG CAGAGTTGGA TACCCCATG TGAAGCCAGG 360  
GCCCAACTGT GGATTTGGAA AAACCTGGCAT TATGTATTAT GGAATCCGTC TCAATAGGAG 420  
TGAAAGATGG GATGCCCTATT GCTACAACCC ACACGCCAAG GAGTGTGGTG GCGTCTTTAC 480  
AGATCCAAAG CAAATTTTTA AATCTCCAGG CTCCCAAAT GAGTACGAAG ATAACCAAAT 540  
CTGCTACTGG CACATTAGAC TCAAGTATGG TCAGCGTATT CACCTGAGTT TTTTAGATTT 600  
TGACCTTGAA GATGACCCAG GTTGCCTGGC TGATTATGTT GAAATATATG ACAGTTACGA 660  
TGATGTCAT GGCTTTGTGG GAAGATACGT TGGAGATGAG CTCCAGATG ACATCATCAG 720  
TACAGGAAT GTCATGACCT TGAAGTTTCT AAGTGATGCT TCAGTGACAG CTGGAGGTTT 780  
CCAAATCAAA TATGTTGCAA TGGATCCTGT ATCCAAATCC AGTCAAGGAA AAAATACAAG 840  
TACTACTTCT ACTGGAATAA AAAACTTTT AGCTGGAAGA TTTAGCCACT TATAAAAAAA 900  
AAAAAAGGA TGATCAAAAC ACACAGTGTT TATGTTGGAA TCTTTTGGAA CTCCTTTGAT 960  
CTCACTGTTA TTATTAACAT TTATTTAITA TTTTCTAAA TGTGAAAGCA ATACATAATT 1020  
TAGGGAATAT TGGAAATAT AGGAACTTT AAACGAGAAA ATGAAACCTC TCATATATCC 1080  
ACTGCATAGA AATAACAAGC GTTAACATTT TCATATTTT TCTTTTCAGT CATTTTCTTA 1140  
TTTGTGGTAT ATGTATATAT GTACCTATAT GTATTTGAT TGGAAATTTT GGAATCCTGC 1200  
TCTATGTACA GTTTTGTATT ATACTTTTAA AATCTTGAA TTTATAAACA TTTCTGAAA 1260  
TCATTGATTA TTCTACAAA ACATGATTTT AAACAGCTGT AAAATATTCT ATGATATGAA 1320  
TGTTTATGCG ATTATTTAAG CCTGTCTCTA TTGTTGGAAT TTCAGGTCAT TTTCATAAAT 1380  
ATTGTTGCAA TAAATATCCT TGAACACACA AAAAAAAA AA

Seq ID NO: 505 Protein sequence  
Protein Accession #: Eos sequence

25  
30

1 11 21 31 41 51  
| | | | |  
MIILYLFL LLWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSQKYLK YAEKAVCEP 60  
EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNCXFGKTGI IDYGIRLNRS 120  
ERWDAYCYNP HAKECGVFT DPKQIFKSPG FPNEDNDQI CYWHIRLKYG QRIHLSFLDF 180  
DLEDDPGCLA DYVEIYDSYD DVEHGFVGRYC GDELPDDIIS TGNVMTLKLFL SDASVTAGGF 240  
QIKYVAMPDV SKSSQGNKTS TTSTGNKNFL AGRFPHL

Seq ID NO: 506 DNA sequence  
Nucleic Acid Accession #: NM\_007115.1  
Coding sequence: 69..902

35  
40  
45  
50  
55  
60

1 11 21 31 41 51  
| | | | |  
GAATTCGCAC TGCTCTGAGA ATTGTGAGC AGCCCCTAAC AGGCTGTAC TTTACTACAA 60  
CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT 120  
GGGGATTCAA GGATGGAATT TTTCAATACT CCATATGGCT TGAACGAGCA GCGGTGTGT 180  
ACCACAGAGA AATCAAGGCT GGCATAATCA AGCTCACCTA CGCAGAAGCT AAGCGGTGT 240  
GTGAATTGTA AGGCGGCCAT CTCGCAACTT ACAAGCAGCT AGAGGCGGCC AGAAAAATTG 300  
GATTTACATG CTGTGCTGCT GGATGGATGG CTAAGGCGAG AGTTGGATAC CCCATTGTGA 360  
AGCCAGGGCC CAACGTATGA TTTGGAAAA CTGGCATTAT TGATTATGGA ATCGGTCTCA 420  
ATAGGAGTGA AAGATGGGAT GCCTATTGCT ACAACCCACA CGCAAGGAG TGTGGTGGCG 480  
TCTTTACAGA TCCAAAGCGA ATTTTAAAT CTCCAGGCTT CCAATATGAG TACGAAGATA 540  
ACCAATCTG CTACTGGCAC ATTAGACTCA AGTATGGTCA GCGTATTCAC CTGAGTTTTT 600  
TAGATTTTGA CTTTGAAGAT GACCCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA 660  
GTTACGATGA TGTCATGGC TTTGTGGGAA GATAGTGGG AGATGAGCTT CCAGATGACA 720  
TCATCAGTAC AGGAATATGC ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG 780  
GAGGTTTCCA AATCAAAATG GTTGCATGG ATCCTGTATC CAAATCCAGT CAAGGAATAA 840  
ATACAAGTAC TACTCTACT GGAATAAAAA ACTTTTATAG TGAAGATTT AGCCACTTAT 900  
AAAAAATAAA AAGGATGATG AAAACACACA GTGTTATGTT TGAATCTTT TGAATCTCT 960  
TTGATCTCAC TGTATTATT AACTTTATT TATTATTTT CTAAATGTGA AAGAAATACA 1020  
TAATTTAGGG AAAATTGGAA AATATAGGAA ACTTTAAAG AGAAAAATGA ACCTCTCATA 1080  
ATCCCACTGC ATAGAAATTA CAAGCTTAA CATTTTCTA TTTTCTCTT TCAGTCATTT 1140  
TTGATTTGT GGTATATGTA TATATGTACC TATATGTATT TGCAATTGAA ATTTTGGAA 1200  
CTGCTCTAT CTTGATTTT GTATTATACT TTTTAAATCT TGAATCTTAT GAACATTTTC 1260  
TGAAATCATT GATTATTCTA CAAAACATG ATTTTAAACA GCTGTAAAT ATTCTATGAT 1320  
ATGAATGTTT TATGCATTAT TTAAGCCTGT CTCTATTGTT GGAATTCAG GTCATTTTCA 1380  
TAAATATTGT TGCAATAAAT ATCCTTCGGA ATTC

Seq ID NO: 507 Protein sequence  
Protein Accession #: NP\_009046.1

65  
70

1 11 21 31 41 51  
| | | | |  
MIILYLFL LLWEDTQGWGF KDGIFHNSIW LERAAGVYHR EARSQKYLK YAEKAVCEP 60  
EGGHLATYKQ LEAARKIGFH VCAAGWMAKG RVGYPIVKPG PNCXFGKTGI IDYGIRLNRS 120  
ERWDAYCYNP HAKECGVFT DPKRIFKSPG FPNEDNDQI CYWHIRLKYG QRIHLSFLDF 180  
DLEDDPGCLA DYVEIYDSYD DVEHGFVGRYC GDELPDDIIS TGNVMTLKLFL SDASVTAGGF 240  
QIKYVAMPDV SKSSQGNKTS TTSTGNKNFL AGRFPHL

75  
80

Seq ID NO: 508 DNA sequence  
Nucleic Acid Accession #: NM\_001044.1  
Coding sequence: 129..1991

85

1 11 21 31 41 51  
| | | | |  
ACCGCTCGG AGCGGGAGGG GAGGCTTGGC GGAACGCTCT CGGCGCCAGG ACTCGCGTGC 60  
AAAGCCCGAG CCGCGGCGGC CAGACCAAGA GGAAGAAGC ACAGAATTCC TCAACTCCCA 120  
GTGTGCCCAT GAGTAAGAGC AAATGCTCCG TGGGACTCCG GTCTTCGCTG GTGGCCCGG 180  
CTAAGGAGCC CAATGCCGTG GCGCGAAGG AGGTGGAGCT CATCTTGTGC AAGGAGCAGA 240  
ACGGAATGCA GCTCACCAGC TCCACCTCA CCAACCCGCG GCAGAGCCCG GTGGAGGCC 300  
AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTCTCCT GTCCGTCAAT GGCTTTGCTG 360  
TGGACCTGGC CAACGCTCGG CGGTTCCTCT ACCTGTGCTA CAAAAATGGT GCGGTGCTCT 420

	TCCTGGTCCC	CTACCTGCTC	TTTCATGGTCA	TTGCTGGGAT	GCCACTTTTC	TACATGGAGC	480
	TGGCCCTCGG	CCAGTTCAAC	AGGGAAGGGG	COGCTGGTGT	CTGGAAGATC	TGCCCCATAC	540
	TGAAGAGTGT	GGGCTTCACG	GTTCATCTCTA	TCTCACTGTA	TGTGGGCTTC	TTCTACAAAG	600
5	TCATCATCGC	CTGGGCGCTG	CACATCTCTT	TCTCTCTCTT	CACCACGGAG	CTCCCTCGGA	660
	TCCACTGCAA	CAACTCTGCG	AACAGCCCA	ACTGCTCGGA	TGCCCATCTCT	GGTGACTCCA	720
	GTGGAGACAG	CTCGGCGCTC	AACGACACTT	TTGGGACCAC	ACCTGCTGCC	GAGTACTTTG	780
	AACGTGGCGT	GCTGCACCTC	CACAGAGGCC	ATGGCATCGA	CGAAGTGGGG	CCTCCGCGGT	840
	GGCAGCTCAC	AGCCTGCGTG	GTGCTGGTCA	TGCTGCTGCT	CTACTTCAGC	CTCTGGAAGG	900
10	GCCTGAAGAC	CTCAGGGAAG	GTGGTATGGA	TCACAGCCAC	CATGCCATAC	GTGGTCTCTA	960
	CTGCCCTGCT	CCTGCTGGGG	GTCCCTCCCT	CTGGAGCCAT	AGAGCGCATC	AGAGCATACC	1020
	TGAGCGTTGA	CTTCTACCGG	CTCTGCGAGG	CGTCTGTTTG	GATTGAOCCG	GCCACCCAGG	1080
	TGTGCTTCTC	CTTGGGCGTG	GGGTTCCGGG	TGCTGATCGC	CTTCTCCAGC	TACAACAAGT	1140
	TCACCAACAA	CTGCTACAGG	GACGCGATTG	TCACCACTCT	CATCAACTCC	CTGACGAGCT	1200
15	TCTCCTCCGG	CTTGTCTGTC	TTCTCCTTCC	TGGGGTACAT	GGCAGAGAAG	CACAGTGTGC	1260
	CCATCGGGGA	CGTGCCCAAG	GACGGGCCAG	GGCTGATCTT	CATCATCTAC	CCGGAAGCCA	1320
	TGCGCAAGCT	CCCTCTGTCC	TCAGCCTGGG	CGGTGCTCTT	CTTCATCATG	CTGCTCAGCC	1380
	TGGGTATCGA	CAGCGCCATG	GGTGTATATG	AGTCACTGAT	CACCGGGCTC	ATCGATGAGT	1440
	TCCAGCTGCT	GCACAGACAC	CGTAGGCTCT	TCAGGCTCTT	CATGCTCCTG	GGGACCTTCC	1500
20	TCTGTCTCCT	GTCTGCGTTC	ACCAACGGTG	GCATCTACGT	CTTCAAGCTC	CTGGACCATT	1560
	TTGACGCGGG	CAGTCCATTC	CTCTTTGGAG	TGCTCATCGA	AGCCATCGGA	GTGGCTTGGT	1620
	TCTATGTGTG	TGGCGAGTTC	AGCGACGACA	TCCAGCAGAT	GACCGGGCAG	CGGCGCAGCC	1680
	TGTACTGGCG	GCTGTGCTGG	AAGCTGGTCA	GCCCCCTGCTT	TCTCCTGTTT	GTGGTCTGTTG	1740
	TCAGCATTTG	GACCTTCAGA	CCCCCCCCCT	ACGGAGCCCTA	CATCTTCCCC	GACTGGGCCA	1800
25	ACGCGCTGGG	CTGGGTCTATC	GCCACATCCT	CCATGGCCAT	GGTGCCCATC	TATGGGCGCT	1860
	ACAAGTTCTG	CAGCCTGCTC	GGGTCTCTTC	GAGAGAAACT	GGCCTAAGCC	ATTGCACCCG	1920
	AGAAGGACCG	TGAGCTGGTG	GACAGAGGGG	AGGTGCGCCA	GTTCAAGCTC	CGCCACTGGC	1980
	TCAGGTGTA	GAGGAGCAG	AGACGAAGAC	CCAGGAAGT	CATCTGCA	TGGGAGAGAC	2040
	ACGAACAAC	CAAGAAATC	TAAGTTTCA	GAGAAAGGAG	GGCAACTTCT	ACTCTTCAAC	2100
30	CTCTACTGAA	AACACAACA	ACAAAGCAGA	AGACTCCTCT	CTTCTGACTG	TTTACACCTT	2160
	TCCGTGCGGG	GAGCGACCTT	CGCGGTGTCT	TGTGTGCTG	TAATAACGAC	GTAGATCTGT	2220
	GCAGCGAGGT	CCAGCCGTTT	GTGTCTCTG	CAGGCGAGAA	AAACGTCTAA	CTTCATGCTG	2280
	TCTGTGTGAG	GCTCCCTCCC	TCCCTGTCTC	CTGCTCCCGG	CTCTGAGGCT	GCCCCAGGGG	2340
	CACCTGTGTT	TCAGGCGGGG	ATCAGATCC	TTGTAGACGC	ACCTGCTGAG	AATCCCCGTTG	2400
35	CTCAGAGTAG	CTTCTAGAC	CATTTACTTT	GCCCCATATTA	AAAAGCCAAAG	TGTCTGCTTT	2460
	GGTTTAGCTG	TGCAGAGGT	GAATGGAGG	AAACCACAAA	TTATGCAAA	GTCTTTTCCC	2520
	GATGCGTGGC	TCCAGCAGA	GGCGTAAAT	TGAGCGTTCA	GTGACACAT	TGCACACACA	2580
	GTCTGTTTCA	AGGCATTGGA	GGATGGGGGT	CCTGGTATGT	CTCACCAGGA	AATTCTGTTT	2640
	ATGTTCTTGC	AGCAGAGAGA	AATAAACTC	CTTGAACCA	GCTCAGGCTA	CTGCCACTCA	2700
40	GGCAGCTGT	GGGTCTTGT	GGTGTAGGGA	ACGCGCTGAG	AGGAGCGTGT	CCATCCCCG	2760
	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	CCTATCCCCG	GACGCATGCA	GGGCCCCCAC	2820
	AGGAGCATGT	CCTATCCCCG	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	ACTACCCAG	2880
	AACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	ACTACCCAG	GACGCATGCA	GGGCCCCCAC	2940
	TGGAGCGTGT	ACTACCCAG	GACGCATGCA	GGGCCCCCAC	AGGAGCGTGT	CCTATCCCCG	3000
45	GACCGAGGCG	ATGCAGGGCC	CCCACAGGAG	CGTGTACTAC	CCCAGGAGCG	ATGCAGGGCC	3060
	CCCACAGGAG	CGTGTACTAC	CCCAGGATGC	ATGCAGGGCC	CCCACAGGAG	CGTGTACTAC	3120
	CCCAGGAGCG	ATGCAGGGCC	CCCATGCGAG	CAGCCTGCGAG	ACCAACACTC	TGCTGGCGCT	3180
	TGAGCGGTGA	CCTCCAGGAA	GGGCCCCCAC	TGGAATTTTA	TTTCTCTCAG	GTGGGTGCCA	3240
	CATCAATAAC	AACAGTTTTT	ATGTTTGGGA	ATGGCTTTTT	AAAATCATAT	TTACCTGTGA	3300
50	ATCAAAACAA	ATTCAAGAA	GCAGTATCCG	CGAGCTCTGT	TGCTGATATT	GCAGTTTTTG	3360
	TTTACAAGAA	TAATTAGCAA	TACTGAGTGA	AGGATGTTGG	CCAAAAGCTG	CTTTCATATG	3420
	CACACTGCCC	TCTGCCACTG	ACAGGAAAGT	GGATGCCATA	GTTTGAATTC	ATGCTCTAAG	3480
	TGCGTGGGCG	TGCCCTACGTG	CTGCCCGAGG	GCAGGGGCGG	TGCAGGGCCA	GTCTGGGCTG	3540
	TCCCTGCAAA	GTGAGCTGGG	GCTCCAGGGA	CTGGAGTGTG	ATGCTGGGTG	GGAGCCGTCA	3600
55	GCCTGTGAAC	TGCCAGGCGA	CTGCAGTTAG	CACAGAGGAT	GGCTTCCCCA	TTGCCCTCTG	3660
	GGGAGGGACA	CAGAGGACGG	CTTCCCCATC	GCCTTCTGGC	CGCTGCAGTC	AGCAGAGAGA	3720
	GCCTGCTCCC	CATTGCTCTC	TGGGGAGGGA	CACAGAGGAG	AGTTTCCCCA	TGCCCTTCTG	3780
	GTGTTTGAAG	ACAGCAGAGA	GAGCGGCTTC	CCCATGCGCT	TCTGGGGAGG	GGCTCCGTGT	3840
	AGCAACCCAG	GTGTTGTCGG	TGCTGTGTTG	CCAATCTCTA	TTGAGCATCG	TGTGGGTCCC	3900
60	TAAGCACAAT	AAAAGACATC	CACAATGGAA	AAAAAAAAG	GAATTC		

Seq ID NO: 509 Protein sequence  
Protein Accession #: NP\_001035.1

65	1	11	21	31	41	51	
	MSKSKCSVGL	MSSVVAFAKE	PNAVGPKEVE	LILVKEONGV	QLTSSTLTNP	RQSFVEAQDR	60
	ETWGGKIDFL	LSVIGFAVDL	ANVWRFPYLC	YKNGGGAFLV	PYLLFMVIAG	MPLFYMELAL	120
	QPFNREGAAG	VMKICPILKG	VGPTVILISL	YVGFYNNVII	AWALHYLPSS	FTTELPWIHC	180
70	NNSWNSPNC	DAHFGDSSGD	SSGLNDTFGT	TPAAEYFERG	VHLHQSHGI	DDLGPFRWQL	240
	TACLVLVIVL	LYPSLNGVKV	TSGRVWVITA	TMPIYVLTAL	LLRGVTLPGA	IDGIRAYLSV	300
	DPYRLCEASV	WIDAATQVCF	SLGVGFGVLI	AFSSYNKFTN	NCYRDAIVTT	SINSLTSFSS	360
	GFVVFSEFLY	MAQKHSVPFG	DVAKDGPGLI	FIIYPEAIAT	LPLSSAWAVV	PFIMLLTLGI	420
	DSAMGMESV	ITGLIDEFQL	LHRHRELFTL	FIVLATFLLS	LPCVTNGGIY	VPTLLDRFAA	480
75	GTSLIFGLVI	EAIGVANFYG	VQGFSDDIQ	MTGQRPSLYW	RLCWKLVSPP	FLLFVVVSI	540
	VIFRPPHYGA	YIFPDWANAL	GWVIATSSMA	MVPIYAAAYF	CSLPGSPREK	LAYALAPEKD	600
	RELVDGRGEVR	QPTLRNWLKV					

Seq ID NO: 510 DNA sequence  
Nucleic Acid Accession #: NM\_001216.1  
Coding sequence: 43..1422

85	1	11	21	31	41	51	
	GCCCGTACAC	ACCGTGTGCT	GGGACACCCC	ACAGTCAGCC	GCAATGGCTCC	CCTGTGCCCC	60
	AGCCCCCTGC	TCCCTCTGTT	GATCCCGGCC	CCTGCTCCAG	GCCTCACTGT	GCAACTGCTG	120
	CTGTCACTGC	TGCTTCTGAT	GCCTGTCCAT	CCCCAGAGGT	TGCCCCGGAT	GCAGGAGGAT	180
	TCCCCCTTGG	GAGGAGGCTC	TCTGGGGGAA	GATGACCCAC	TGGGGGAGGA	GGATCTGCC	240

AGTGAAGAGG ATTCAACCAG AGAGGAGGAT CCACCCGGAG AGGAGGATCT ACCTGGAGAG 300  
 GAGGATCTAC CTGAGAGAGA GGATCTACCT GAAGTTAAGC CTAATCAGA AGAAGAGGGC 360  
 TCCTGAAGT TAGAGGATCT ACCTACTGTT GAGGCTCCTG GAGATCCTCA AGAAGAGGGC 420  
 AATAATGCCC ACAGGGACAA AGAAGGGGAT GACCAGAGTC ATTGGCGCTA TGGAGGGGAC 480  
 CGGCGCTGGC CCGGGGTGTC CCGAGCCTGC GCGGGCGCTT TCCAGTCCCC GGTGGATATC 540  
 CGGCGCGAGC TGCGCGCTTT CTGCGCGGCC CTGCGCGGCC TGGAACTCCT GGGCTTCCAG 600  
 CTCCCGCGGC TCACAGAACT GCGCTGCGC AACAATGGCC ACAGTGTGCA ACTGACCCCTG 660  
 CCTCTGGGCG TAGAGATGGC TCTGGGTCCC GGGCGGGAGT ACCGGGCTCT GCAGCTGCAT 720  
 CTGCACTGGG GGGCTGCAGG TCGTCCGGGC TCGGAGCACA CTGTGGAAGG CCACCGTTTC 780  
 CCTGCGGAGA TCCAGTGGT TCACTCAGC ACCGCTTTG CCAGAGTTGA CGAGGCTTTC 840  
 GGGCGCGGCG GAGGCGCTTC CGTGTTGGCC GCCTTTCTGG AGGAGGGGCC GGAAGAAAAC 900  
 AGTGCCTATG AGCAGTTGCT GTCTCGCTTG GAAGAAATCG CTAGGAAGG CTCAGAGACT 960  
 CAGGTCCCAG GACTGGACAT ATCTGCACTC CTGCGCTCTG ACTTCAGCG CTACTTCCAA 1020  
 TATGAGGGGT CTCTGACTAC ACCGCGCTGT GCCCAGGGTG TCACTGGAC TGTGTTTAA 1080  
 CAGACAGTGA TGCTGAGTGC TAAGCAGCTC CACACCTCT CTGACACCT GTGGGGACCT 1140  
 GGTGACTCTC GGTACAGCT GAACTTCCGA GCGAGCGAGC CTTTGAATGG GCGAGTGATT 1200  
 GAGGCTCTCT TCCTGCTGG AGTGGACAGC AGTCTCGGG CTGCTGAGCC AGTCCAGCTG 1260  
 AATTCTCTGC TGGCTGCTGG TGACATCCTA GCGCTGTTT TTGGCTCCT TTTTGTCTG 1320  
 ACCAGCGTGG CGTTCCTTGT GCAGATGAGA AGGCAGCACA GAAGGGGAAC CAAAGGGGGT 1380  
 GTGAGTACCC GCCCAGCAGA GGTAGCCGAG ACTGGAGCCT AGAGGCTGGA TCTTGGAGAA 1440  
 TGTGAGAGAG CAGCCAGAGG CATCTGAGG GAGAGCGGTA ACTGTCCTGT CCTGCTCAT 1500  
 ATGCCACTTC CTTTAACTG CCAAGAAATT TTTTAAATA AATATTTATA AT

Seq ID NO: 511 Protein sequence  
Protein Accession #: NP\_001207.1

1 11 21 31 41 51  
 MAPLCPSPWL PLLIPAPAPG LTVQLLSLL LLMPVHFQRL PRMQEDSPLG GSSGSEDPL 60  
 GEEDLPSEED SPREDPPGE EDLPGEEDLP GEEDLPVKK KSEEGSLKL EDLPTEAPG 120  
 DPQEPNNNAH RDKEGDDQSH WRYGDPFPP RVSPACAGRP RSPVDIRPQL AAFPALRPL 180  
 ELLGFLQPL PELRLNNGH SVQLTLPPLG EMALGPRHY RALQLHLHWG AAGRPGSEHT 240  
 VEGHFPFAEI HVHLLSTAF RVDALGRFG GLAVLAFL EGPENSAEY QLSRLLEIA 300  
 EGGSETQVPG LDISALLPSD FSRYPQYEGS LTPPACQGV IWTVPNTVM LSAQLHTLS 360  
 DTLWNGPDSR LQLNFRATQP LNRVIEASF PAGVDSPPRA AEPVQLNSCL AAGDILALVF 420  
 GLLFAVTSVA FLVQMRQR RGTGKGVSYR PAEVAETGA

Seq ID NO: 512 DNA sequence  
Nucleic Acid Accession #: Bos sequence  
Coding sequence: 1..3978

1 11 21 31 41 51  
 ATGTTGGGTG AAGGACCTTA CTTATCTCA GATCTGGACC AGCGAGGCGG GCGGAGATCC 60  
 TTGTCAGAAA GATATGACCC CAGCCTGAAG ACCATGATCC CAGTGCAGCC CTGTGCAAGG 120  
 TTAGACCCCA ACCCGGTGGA TGATGCGGG CTACTCTCCT TCGCCACATT TTCCTGGCTC 180  
 ACCGCGGTGA TGTGAAAGG CTACCGGCAA AGGCTGACCG TAGACACCTT GCCCCATTG 240  
 TCGACTATG ACTCATCTGA CACCAATGCC AAAAGATTTC GAGTCTTTG GGAATGAAG 300  
 GTAGCAAGGG TGGGTCTTGA GAAGGCTCT CTGAGCCACG TGGTGTGAA ATTCCAGAG 360  
 ACACGCGTGT TGATGGACAT CGTGGCCAAC ATCCTGTGCA TCATCATGCG AGCCATAGG 420  
 CGACAGTTT CTATTCAACA AATCCTCCAG CAGACTGAGA GGACCTCTGG GAAAGTCTGG 480  
 GTTGGCATTG GACTGTGAT AGCCCTTTT GCCACCGAGT TTACCAAGT CTCTCTTTGG 540  
 GCCTTTGCTT GCGCCATCAA CTACCGCAGC GCCATCGGT TGAAGGTGGC GCTCTCCACC 600  
 TTGTTTGGT AAAACCTAGT GTCTTCAAG ACATGACCC ACATCTCTGT TGGCGAGGTG 660  
 CTCATATATC TGTCAGTGA TAGCTATTCT TTGTTGAAG CTGCTTGTG TTGTCTTTG 720  
 CCAGCCACCA TCCCGATCCT AATGGTCTTT TGTGCGGCGT ACGCTTTT CTCTCTGGG 780  
 CCCACAGCTC TCATCGGGAT ATCAGTGTAT GTCATATTCA TACCGTCCA GATGTTTATG 840  
 GCCAAGCTCA ATTACGCTTT CCGAAGGTCA GCAATTTTGG TGACAGACAA GCGAGTTGAG 900  
 ACAATGAATG AGTTTCTGAC CTGCATCAGG CTGATCAAAA TGATAGCTG GGAAGAAATCT 960  
 TTTACCAACA CTATCCAGA TATAAGAAG AGGGAAGAA AATTACTGGA AAAAGCTGGA 1020  
 TTGTCCAAA GTGAAACTC TGCCCTGGCC CCCATCGGT CCACCATAGC CATCGTGTG 1080  
 ACATATCTCT GCCACATCCT CTTGAGACGC AAATCAGCC CACCGTGGC ATTTAGTGTG 1140  
 ATTGCCATGT TTAATGTAAT GAAGTTTTC ATTGCAATCT TGCCCTTCT CATCAAAGCA 1200  
 ATGCTGAAG CGAATGTCT TCTAAGGAGA ATGAAGAAA TTCTCATAGA TAAAGCCCTC 1260  
 CCATCTTACA TCACCCACCC AGAAGACCCA GATACTGTCT TGCTTTAGC AAATGCCACC 1320  
 TTGACATGGG AGCATGAAGC CAGCAGGAAA AGTACCCCAA AGAATTTGCA GAACAGAAA 1380  
 AGGCATTATG GCAAGAAACA GAGGTGAGG GCATACAGT AGAGGAGTCC ACCAGCCAG 1440  
 GGAGCCACTG GCCCAGAGGA GCAAGTGAC AGCCTCAAAT CGTTCTGCA CAGCATAAGC 1500  
 TTTGTGGTGA GAAAGTTATG TCGTTATCCC GAAGCCGAGC TCCTGGCTTG GAGGTGGCCA 1560  
 GCACTGTTTG TTGGGAGAAT CATCAGAGGA TACAGGCTTC ATGGATTTTC TGCTAAAGAC 1620  
 AAGATGAAT CTAGAAGGCT TCTTACTTGG CCCAAGAAAG TGATAGGAC TCAAGGGCA 1680  
 GCCAATATCC TGGGGAAGAT CTTGGGAATA TGTGGGAATG TGGGAAGTGG AAGAGCTCC 1740  
 CTCTTGCAG CTCTCTTAGG ACAGATGCAG CTGCAGAAAG GGGTGGTGGC AGTCAATGGA 1800  
 ACTTGGGCTC ACCTTTTACA GCAGGATGG ATCTTTTATG GAAATGTGAG AGAAACATA 1860  
 CTCTTTGGAG AAAAGTATGA TCACCAAGG TATCAGCACA CAGTCCGCT CTGTGGCCTC 1920  
 CAGAGGACC TGAGCAACCT CCCCTATGGA GACCTGACTG AGATTGGGGA GCGGGGCTC 1980  
 AACCTCTCTG GGGGCGAGG GCAGAGGATT AGCCTGGCCC GCGCTGTCTA CTCGACCGT 2040  
 CAGCTCTACC TGCTGAGAGA CCCCTGTGCG GCCGTGGAGC CCCAGTGGG GAAGCAGCTC 2100  
 TTTGAGGAGT GCATTAAAGG GAGCTCAGG GGAAGACAG TGTCTCTGTT GACCCACAG 2160  
 CTACAGTCTT TAGAGTCTTG TGATGAAGTT ATTTTATTAG AAGATGGAGA GATTGTGAA 2220  
 AAGGGAACCC ACAAGAGATT AATGAGGAG AGAGGGCGCT ATGCAAACT GATTCACAAC 2280  
 CTGCGAGGAT TGCAGTTCAA GATCTCTGAA CACCTTTACA ATGCAGCAAT GGTGGAAGCC 2340  
 TTCAGGAGA GCCCTGCTGA GAGAGAGGAA GATGCTGTA TAATCGGTA CCTCTTTCT 2400  
 CTCTCACTG TGTCTCTCTT CCTCTGATG ATTGGCAGCG CTGCTCTCAG CAAGTGGTG 2460  
 CTGGGCTCTT GGTGACAAA GGGCTCACGG ATGACCTGTG GCGCCAGGG CAACAGGACC 2520  
 ATGTGTGAGG TCGCGCGGCT GCTGGCAGC ATCGGTGAG ATGTGTACCA GTGGGTGTAC 2580  
 ACTGCAAGCA TGGTGTTCAT GCTGGTGTG GCGCTCACA AAGGCTTCT CTTCACCAAG 2640

ACCACACTGA TGGCATCTCT CTCTCTGCAT GACACGGTGT TTGATAAGAT CTTAAAGAGC 2700  
 CCAATGAGTT TCTTTGACAC GACTCCCACT GGCAGGCTAA TGAACCGTTT TTCCAAGGAT 2760  
 ATGGACGAGC TGGATGTGAG GCTCGCGTTT CACGACAGAA ACTTTCTGCA GCAGTTTITT 2820  
 ATGGTGTGTG TTATTCTCGT GATCTTGGCT GCTGTGTTTC CTGCTGTCCCT TTATGTCGTG 2880  
 GCCAGCCTTG CTGTAGGCTT CTTCAATCTG TTACGCATT TCCACAGAGG AGTCCAGGAG 2940  
 CTCAGAAGG TGGAGAATGT CAGCGGTCA CCTTGGTTCA CCCACATCAC CTCCTCCATG 3000  
 CAGGCGCTGG GCATCATTC CCGCTATGGC AAGAAGGAGA GCTGCATCAC CTATACTTCA 3060  
 TCCAAGGCC TGTCAATGTC ATACATCATC CAGCTGAGCG GACTGTCTCA AGTGTGTGTG 3120  
 CGAAGCGGAA CAGAGACGCA AGCCAAATTC ACCTCGGTGG AGCTGCTCAG GGAATACATT 3180  
 TCGACCTGTG TTCTGAATG CACTCATCCC CTCAAAGTGG GGACCTGTCC CAAGGACTGG 3240  
 CCCAGCTGTG GGGAGATCAC CTTGAGAGAC TATCAGATGA GATACAGAGA CAACACCCCC 3300  
 CTTGTTCTCG ACAGCCTGAA CTTGAACATA CAAAGTGGGC AGACAGTCGG GATTGTTGGA 3360  
 AGAACAGTT CCGGAAAGTC ATCGTTAGGA ATGCTTTTGT TTGCTCTGGT GGAGCCAGCC 3420  
 AGTGGCACAA TCTTTATTTA TGAGGTGGAT ATCTGCATT TCAGCTTGA AGACCTCAGA 3480  
 ACCAAGCTGA CTGTGATCCC ACAGGATCCT GTCTCTGTTG TAGGTACAGT AAGGTACAAC 3540  
 TTGATCCCT TTGATGATG CACCGATGAG ATGCTCTGGC AGGTTCTGGA GAGAATCATT 3600  
 ATGAGAGACA CAATAATGAA ACTCCAGAA AAATTACAGG CAGAAGTCAC AGAAAATGGA 3660  
 GAAAACCTCT CAGTAGGGGA ACGTCAGCTG CTTTGTGTGG CCCGAGCTCT TCTCCGTAAT 3720  
 TCAAGATACA TTCTCTTGA TGAAGCCACC GCCTCTATGG ACTCCAAGAC TGACACCCGTG 3780  
 GTTCAGACA CCATCAAGA TGCCTTCAAG GGCTGCACTG TGCTGACCAT CGCCACCGGC 3840  
 CTCACACAG TTCTCAACTG GCATCACGTC CTGGTTATGG AAAATGGGAA GGTGATTGAG 3900  
 TTTGACAAGC CTGAAGTCCT TGCAGAGAAG CCAGATTCTG CATTTCGGAT GTTACTAGCA 3960  
 GCAGAAGTCA GATTGTAG

Seq ID NO: 513 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MVGEGPYLIS DLDQRGRRRS FAERYDPSLK TMIPVVRPCAR LAPNPVDDAG LLSFATFWSL 60  
 TPVMVKGYRQ RLTVDTLPL STYDSSDTNA KRFRVLWDEE VARVGPKEAS LSHVVVKFQR 120  
 TRVLMDIVAN ILCLIMAAIG PTVLHQLIQ QTERTSGKVV VIGLCLIALF ATEFTKVFFW 180  
 ALAWAINYRT AIRLKVALLST LVFENLVSPK TLTHISVGEV LNILSSDSYS LFEAALFCPL 240  
 PATIPILMVF CAAYAPFILG PTALIGISVY VIPFVQMPM AKLNSAFRRS AILVTDKRVQ 300  
 TMNEFLTICR LIKMYAWEKS PTNTIQDIRR RERKLEKAG PVQSGNSALA PIVSTIAIVL 360  
 TLSCHILLRR KLTAPVAFSV IAFNVKMFPS IAILPFSIKA MAEANVSLRR MKKILIDKSP 420  
 PSYITQPEPD DTULLANAT LTWEHRSRK STPKLQNKQ RHLCKQRSE AYSESPPAK 480  
 GATGPEQSD SLKSVLHSLF FVVRKLCRYF EAQLLAWRNP AVFVGRIRG YRPHGFSAKD 540  
 XDESRLRLTW PQEVDRTQRA AKYLKILGI CGNVGSGKSS LLAALLGQMQ LQKGVVAVNG 600  
 TLAYVSQQA IFHGNVRENI LFGEKYDHQR YQHTVRVCGL QKDLNLPFY DLTEIGERGL 660  
 NLSGGQRQRI SLARAVYSR QLYLLDDPLS AVDAHVGKHV FEBCIKKTLR GKTVVLVTHQ 720  
 LQPLESCDEV ILLEDEGEICB KGTHELMEE RGRYAKLIHN LRGLQPKDPE HLYNAAMVEA 780  
 FKESPAEREE DAGIIGYLLS LFTVFLFLM IGSAAFSNNW LGLWLDKGRS MTCGPQGNRT 840  
 MCEVGAUPLD IGRVYQWVY TASVMFLVF GVTKGFVFTK TTLMASSSLH DTVFDKILKS 900  
 MCEFFDTTPT GRLMNRFSD MDELVRLPF HAENFLQPPF MVVFILVILA AVFPAVLLVV 960  
 ASLAVGFFIL RLIFHRGVQE LKKVENVRS PWFTHITSSM QGLGIHAYG KKECITYTS 1020  
 SKGLSLSYII QLSGLLQVCV RTGTETQAKF TSVELLREYI STCVPECTEP LKVGTCPKDW 1080  
 PSCGEITFRD QCMRYRDNTP LVLDLNLNI QSGQTVGIVG RTGSGKSLG MALFRLVEPA 1140  
 SGTIFIDEVD ICILSLDLR TKLTVIPQDP VLFVGTIVRY LDPFESHTE MLNQVLERTF 1200  
 MRDTIMKLEP KLQAEVTENG ENFSVGERQL LCVARALLRN SKIILLDEAT ASMDSKDTL 1260  
 VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENGKIVIE FDKPEVLAKK PDSAFAMLLA 1320  
 AEVRL

Seq ID NO: 514 DNA sequence  
 Nucleic Acid Accession #: Z31560  
 Coding sequence: 1-966

1 11 21 31 41 51  
 CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCGGGG CCGCAGCAA 60  
 ACTTCGGGGG GCGGCGGGCG CAACTCCACC GCGGCGGGCG CCGGCGGCAA CCAGAAAAAC 120  
 AGCCCGGACC CGGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CCGGCAGCGG 180  
 CGCAAGATGG CCCAGGAGAA CCCCAGATG CACAACCTCG AGATCAGCAA GCGCCTGGGC 240  
 GCGGAGTGA AACTTTTGTG GAGACGGAG AAGCGGCGGT TCATCGACGA GGTAAAGCGG 300  
 CTGCGAGCGC TGACATGAA GGAGCACCGG GATTATAAT ACCGGCCCCG CCGGAAAACC 360  
 AAGACGCTCA TGAAGAAGGA TAAGTACAG CTGCCCGCGG GGCTGTCTGC CCCCGCGGCG 420  
 AATAGCATGG CGAGCGGGGT CGGGGTGGGC CCGGCGCTGG GCGCGGGGT GAACAGCGC 480  
 ATGACAGTT ACGCGCACAT GAACGGCTGG AGCAACGGCA GCTACAGCAT GATGCAGGAC 540  
 CAGCTGGGT ACCCGCAGCA CCGGGGCTCT AATGCGCAGG GCGCAGCGCA GATGCAGCCC 600  
 ATGCACCGCT ACGACGTGAG CGCCCTGCAG TACAACCTCA TGACCACTCT GCAGACCTAC 660  
 ATGAACGGCT CGCCACCTA CAGCATGTCC TACTCGCAGC AGGGCACCCC TGGCATGGCT 720  
 CTTGGCTCCA TGGGTTTGGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCCT TGTGTTTACC 780  
 TCTTCTCCC ACTTCAGGGC GCGCTGCCAG CCGGGGACC TCGGGACAT GATCAGCATG 840  
 TATCTCCCC GCGCGGAGGT GCGGGAACCC GCGGCCCCCA GCAGACTTCA CATGTCCCAG 900  
 CACTACAGA GCGGCCCCGT GCGGCGCAGC GCGATTAAAG GCACACTGCC CCTCTCACAC 960  
 ATGTGAGGCG CGGACAGCGA ACTGGAGGGG GGAGAAATTT TCAAGAAAAA ACGAGGGAAA 1020  
 TGGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCCGG TACGCTCAAA 1080  
 AAAAA

Seq ID NO: 515 Protein sequence  
 Protein Accession #: CAA83435

1 11 21 31 41 51  
 HSARMYMMME TELKPPGPQQ TSGGGGGNST AAAAGGNQKV SPDRVKRPMN AFMVWSRGQR 60  
 RKMAQENPKM HNSEISKRLG AEWKLLSETE KRFIDEAKR LRALHMKRHP DYKYRPRRKT 120  
 KTLMKDKYIT LPGLLAPGG NSMASGVGVG AGLGAGVNQR MDSYAHMNGW SNGSYSMMQD 180

QLGYPHQHPGL NARGAAQMQP MHRYDVSALQ YNSMTSSQTY MNGSPITYSMS YSQOQTPGMA 240  
 LGSMSGVVKKS EASSSPFVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPPEP AAPSRLLHMSQ 300  
 HYQSGFVPGT AINGTLPLSH M

Seq ID NO: 516 DNA sequence  
 Nucleic Acid Accession #: U91618  
 Coding sequence: 29..541

1 11 21 31 41 51  
 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAAATCC AGCTTGTATG 60  
 CATGCTACTC CTGGCTTTCA GCTCCTGGAG TCTGTGCTCA GATTTCAGAG AGGAAATGAA 120  
 AGCATTAGAA GCAGATTTC TGAACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180  
 TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAATT TGAACAGCCC 240  
 AGCTGAGGAA ACAGAGAGAG TTCTATGAAGA GGAGCTTGTG GCAAGAAGGA AACTTCCTAC 300  
 TGCTTTAGAT GGCCTTTAGT TGGAAAGCAAT GTTGACAATA TACCAGCTCC ACAAAATCTG 360  
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTTG ATACTGGAAG 420  
 TGACAAAAT GGAAGGAGAG AAGTCATAAA GAGAAAATTT CCTTATATTC TGAACGCGCA 480  
 GCTGTATGAG AATAAACCCA GAAGACCCCTA CATACTCAAA AGAGATTCTT ACTATTACTG 540  
 AGAGAATAAA TCATTTATTT ACATGTGATT GTGATTTCAT ATCCCTTAAT TAAATATCAA 600  
 ATTATATTGT TGTGAAAATG TGACAAACAC ACTTATCTGT CTCTCTACA ATTTGTGGTT 660  
 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720  
 TCTTCAAAA AAAAAAAAAA AAATGGGGCC GCAATT

Seq ID NO: 517 Protein sequence  
 Protein Accession #: AAB50564

1 11 21 31 41 51  
 MMAGMKIQLV CMLLLAFSSN SLCSDEEEM KALEADFLTN MHTSKISKAH VPSWRMTLLN 60  
 VCSLVNINLNS PAEETGEVHE EELVARRKLP TALDGFSLA MLTIYQLKHI CHSRAFQHWE 120  
 LIQEDILDTG NDRNGKEEVI KRKIPYILKR QLYENKPRRP YILKRDSTYY

Seq ID NO: 518 DNA sequence  
 Nucleic Acid Accession #: NM\_006536.2  
 Coding sequence: 109..2940

1 11 21 31 41 51  
 ACCTAAAACC TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60  
 ATGTATGCAG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120  
 AGCATTGCAG GTCTTATTGT CAACCTGAAG TTTGTGACTC TCCTGGTTCG CTTAAGTTCA 180  
 GAACTCCATC TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240  
 ATTGCAATTA ATCCTCAGGT ACCTGAGAAT CAGAACCTCA TCTCAACAT TAAGGAAATG 300  
 ATAACTGAAG CTTCATTTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTGAGAAAT 360  
 ATAAAGATT TAATACCTGC CACATGGAAG GCTAATAATA ACAGCAAAAT AAAAAAGAA 420  
 TACATGAAA AGGCAAAATGT CATAGTGAAT GACTGGTATG GGGCAGATGG AGATGATCCA 480  
 TCACACCTAC AATACAGAGG GTGTGGAAGA GAGGGAATAT ACATTCATTT CACACCTAAT 540  
 TTCCTACTGA ATGATAACTT AACAGCTGGC TACGATCAC GAGGCGAGT GTTTGTCCAT 600  
 GAATGGGGCC ACCTCCGTGT GGTGTGTGTC GATGAGTATA ACAATGACAA ACCTTTCTAC 660  
 ATAAATGGGC AAAATCAAA TAAAGTGACA AGGTGTTTAT CTGACATCAC AGGCATTTTT 720  
 GTGTGTGAAA AAGGTCTCTG CCCCAGGAAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780  
 GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840  
 AGTTTATCTT CTGTGGTTGA ATTTGTGAAT GCAAGTACCC ACAACCAAGA AGCACCACAA 900  
 CTACAGAAC AGATGTGCAG CCTCAGAAGT GCATGGGATG TAATCACAGA CTCTGCTGAC 960  
 TTTACACACA GCTTTCCCAT GAATGGGACT GAGCTTCCAC CTCCTCCAC ATTCTCGCTT 1020  
 GTACAGGCTG GTGACAAAGT GGTCTGTGTT GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080  
 GCTGACAGAC TCCTTCAACT ACAACAAGCC GCAGAAITTT ATTTGATGCA GATTGTTGAA 1140  
 ATTATACCTC TGTGGGATC TGCCAGTTTC GACAGCAAGG GAGAGATCAG AGCCAGCTA 1200  
 CACCAATTA ACAGCAATGA TGATCGAAG TGTGCTGTT CATATCTGCC CACCACTGTA 1260  
 TCAGCTAAAA CAGACATCAG CATTTGTTCA GGGCTTAAGA AAGGATTGGA GGTGGTTGAA 1320  
 AAACCTGAAT GAAAGCTTGA TGGCTCTGTG ATGATATTAG TGACAGCGG AGATGATAAG 1380  
 CTCTGTGGCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCATTGCC 1440  
 CTGGGTTCAT CTGACGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500  
 TTCTTTGTTT CAGATATATC AAACCTCAAT AGCATGATTG ATGCTTTCAG TAGAATTTCC 1560  
 TCTGGAATG GAGACATTTT CCAGCAACAT ATTCAGCTTG AAGTACAGG TGAATAATGC 1620  
 AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680  
 ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCTGAGA TTATATTATT TGATCCTGAT 1740  
 GGAAGAAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTTCGGAC AGCTAGTCTT 1800  
 TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTGAACAA TACCATCAT 1860  
 TCTCTGCAAG CCTGAAAGT GACAGTGACC TCTGCGCCCT CCAACTCAGC TGTGCCCCA 1920  
 GCCACTGTGG AAGCCTTTGT GGAAGAGAGC AGCCTCCATT TTCCTCATCC TGTGATGATT 1980  
 TATGCAAATG TGAACAGGG ATTTTATCCC ATTTCTAATG CCAGTGTAC TGCCACAGTT 2040  
 GAGGCAGAGA CTGAGATGCC TGTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100  
 GTTATAAAAA ATGATGGAAT TTACTCGAGG TATTTTCTCT CCTTGTCTGC AAATGGTAGA 2160  
 TATAGCTTGA AAGTGCAATG CAATCACTCT CCCAGCATAA GCACCCAGC CCACTCTATT 2220  
 CCAGGGAGTC ATGCTATGTA TGTACCAGGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280  
 GCTCCAAGGA ATTCAGTAGG CAGAAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340  
 AGCTCAGGAG GCTCCTTTTC AGTGTCTGGA GTTCCAGCTG GCCCCACCC TGATGTGTTT 2400  
 CCACCATGCA AATATTATGA CCTGGAAGT GTAAGAGTAG AAGAGGAATT GACCCTATCT 2460  
 TGACAGCAC CTGAGAGAAG CTTTGTATCAG GGCCAGGCTA CAGCTATGA AATAAGAAATG 2520  
 AGTAAAGTC TCAGAAATAT CCAAGATGAC TTTAACAATG CTATTTTATG AATATACATA 2580  
 AAGCGAATC TCACAGAACG TGGCATCAGG GAGATATTTA CGTCTCACC CCAGATTTC 2640  
 AGCAATGGAC CTGAACATCA GCCAATGGA GAAACACATG AAAGCCACAG AATTTATGTT 2700  
 GCAATACGAG CAATGCTAG GAACTCCTTA CAGTCTGCTG TATCTAACAT TGCCAGGCG 2760  
 CCTCTGTTTA TTCCCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAAA 2820  
 GGAGTTTAA CAGCAATGGG TTTGATAGGA ATCATTTGCC TTATTATAGT TGTACACAT 2880

CATACTTTAA GCAGGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATAA 2940  
 ATAAATATCC AAAGTGTCTT CCTTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAAA 3000  
 CATACTAACA AAGTCAAATT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTGACA 3060  
 ATACAGATAA GATTTTACCA TGGTAGATCA ACAATTCITT TTGGGGGTAG ATTAGAAAAA 3120  
 CCTTACACTT TGGCTATGAA CAAATAATAA AAATATTCTT TTAAGTAAT GTCTTTAAAG 3180  
 GCAAAGGGAA GGGTAAAGTC GACCCAGTGT CAAGGAAAGT TTGTTTTATT GAGGTGGAAG 3240  
 AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAACTGTCTG TGTGAAGCAA 3300  
 TCATTAGATT ACTTTGATTA ATTTTCTTTC TCTCCTTATC TGTGCAGTAC AGGTGTGCTG 3360  
 TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAAG CTCCTTACCT 3420  
 CTGTGCTATT TGTATATAT ATTTTCAGATG ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480  
 TTTCACTGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540  
 TTTATGACAA AGGTCTATTG AATTTATTGG TGTGTAAGTT TCTACTCCCA TCAAAGCAGC 3600  
 TTTCTAAGTT TATTGCTTGG GTTATTATG GAATGATAGT TATAGCCCCN TATAATGCCT 3660  
 TACCTAGGAA A

Seq ID NO: 519 Protein sequence  
 Protein Accession #: NP\_006527.1

1 11 21 31 41 51  
 MTQRSIAGPI CNLKFVTLIV ALSSELPPILG AGVQLQDNGY NGLLIAINPO VPENQNLIEN 60  
 IKEMITEASF YLFNATKRRV FPRNIKILIP ATWKANNNSK IKQBSYERAN VIVTDWYGAH 120  
 GDDPYTLQYR GCGEGKYIH FTPNFLNDN LTAGYSGRGR VVHEWAHLR WGVFDEYNND 180  
 KPFYINGNQ IKVTRCSSDI TGIFVCEKGP CPQENCIISK LPKEGCTFIY NSTQNATASI 240  
 MFMQSLSSVV EFCNASTENQ EAPNLQNMOC SLRSANDVIT DSADPHHSFP MNGTELPPPP 300  
 TFSLVQAGDK WVCLVLDVSS KMAEADRLLO LQQADEFYLM QIVEIHTFVG IASFDKSGEI 360  
 RAQLHQINSN DDKRLVSVYL PTTVSAKTDI SICSLKKGF EVVEKLNGKA YGSVMILVTS 420  
 GDDKLLGNCL FTVLSSGSTI HSIALGSSAA PNLEELSLRT GGLKFFVPDI SNSNSMIDAF 480  
 SRISSTGDI FQTHQLEST GENVKPHHQL KNTVTVDNVT GNDTMTFLVT QASGPPEIIL 540  
 FDPDGRKYYT NNPITNLTFR TASLWIPGTA KPGHWITLNL NTHSLQALK VVTVSASNS 600  
 AVPPATVEAF VERDSLHFPV PVMIVANVKQ GFYFILNATV TATVEPSTGD FVTLRLDDG 660  
 AGADVIKNDG IYSRYFFSFA ANGRVSLKVH VNHSPSISTP AHSIPGSHAM YVPGYTANGN 720  
 IQMNAPRKSV GRNEEBERKWG FSRVSSGSGF SVLGVPAGPH PDVFPCKII DLEAVKVEE 780  
 LTLSTWAPGE DFDQQAQTSY EIRMSKSLQN IQDDPNNAIL VNTSKRNPQQ AGIREIFTFS 840  
 PQISTNGPEH QPNGETHESH RIYVAIRAMD RNSLQSAVSN IAQAPLFIPP NSDPVPARDY 900  
 LILKGVLTAM GLIGIICLII VVTHHTLSRK KRADKKENGTL KLL

Seq ID NO: 520 DNA sequence  
 Nucleic Acid Accession #: NM\_000228.1  
 Coding sequence: 82..3600

1 11 21 31 41 51  
 GCTTTCAGGC GATCTGAGAA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTTCTGG 60  
 GGATCACCCC ATTGGCTGAA GATGAGACCA TTCTTCCTCT TGTGTTTTCG CCTGCTCTGC 120  
 CTCTGCGATG CCAACAAGCG CTGCTCCGCT GGGGCTGCTG ATCCACCTGT TGGGGACCTG 180  
 CTGTGTGGGA GGACCCGGTT TCTCCGAGCT TCATCTACCT GTGGACTGAC CAAGCCTGAG 240  
 ACCTACTGCA CCGAGTATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCCAGGCAG 300  
 CCTCACAACT ACTACAGTCA CCGAGTAGAG AATGTGGCTT CATCTCCGGG CCCCATGCGC 360  
 TGGTGGCAGT CCGAGAAATGA TGTGAACCTT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420  
 TTCCAGCTTC AAGAAGTCAAT GATGGAGTTC CAGGGGCCCA TGCCCGCCGG CATGCTGATT 480  
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540  
 ACCCTCAACT TCCTCGGGT CCGCCAGGCT CGGCTCAGA GCTGGCAGGA TGTTCGGTGC 600  
 CAGTCCCTGC CTCAGAGGCC TAATGCACGC CTAAATGGGG GGAAGGTCCA ACTTAACTTT 660  
 ATGGATTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAA AAATTCAGA GGTGGGGGAG 720  
 ATCAAACT TGAGAGTCAA TTTCACAGG CTGGCCCTGT TGCCCCAAG GGGCTACCAC 780  
 CCTCCACGCG CTAATATGTC TGTGTCCAG CTCCGTCTGC AGGGGAGCTG CTCTGTGTCAC 840  
 GGCCATGCTG CTGCTGCGC ACCCAAGCCT GGGGCTCTG CAGGCCCTCT CACCGCTGTG 900  
 CAGGTCCACG ATGTCTGTGT CTGCCAGCAC AACACTGCGG GCCCAATTG TGAGCGCTGT 960  
 GCACCTTCT ACAACACCG GCCCTGGAGA CCGGCGGAGG GCCAGGACGC CCATGAATGC 1020  
 CAAAGGTGCG ATGCAATGG GCATCAGAG ACATGTCACT TTGACCCCGC TGTGTTTGC 1080  
 GCGACCCAGG GGCATATGG AGGTGTGTGT GACAATTGCC GGGACACAC CGAAGGCAAG 1140  
 AACTGTGAGC GGTGTCACT GCACATTTTC CGGAACCGGC GCCCGGGAGC TTCCATTGAG 1200  
 GAGACCTGCA TCTCTGCGA GTGTGATCGG GATGGGCGAG TGCCAGGGGC TCCTGTGAC 1260  
 CCAAGTACCG CGAGTGTGT GTGCAAGGAG CATGTGCAGG GAGAGCGCTG TGACCTATGC 1320  
 AAGCCGGGCT TCACTGGACT CACCTACGCC AACCCGAGG GCTGCCACCG CTGTGACTGC 1380  
 AACATCCTGG GTTCCCGGAG GCACATGCGG TGTGACAGG AGAGTGGGCG CTGCTTTGT 1440  
 CTGCCCAACG TGTGGGTCC CAAATGTGAC CAGTGTGCTC CCTACCACTG GAAGCTGGCC 1500  
 AGTGGCCAGG CTGTGAAACC GTGTGCTGCG GACCCGACA ACTCCCTCA GCCCAGAGTG 1560  
 CAACCACTTC ACAGGCGAGT GCCCTGTGCG GAAGGCTTTG GTGCTGTGAT GTGCAGCGCT 1620  
 GCAGCATCC GCCAGTGTCC AGACCGGACC TATGGAGACG TGGCCACAGG ATGCCGAGCC 1680  
 TGTGACTGTG ATTTCCGGGG AACAGAGGCG CCGGCTGCGG ACAAGGCATC AGGCCGCTGC 1740  
 CTCTGCCGCC CTGGCTTGAC CGGGCCCCGC TGTGACAGT GCCAGCGAGG CTAAGTCAAT 1800  
 CGTACCCGCG TGTGCTGGC CTGCCACCTT TGCTTCCAGA CCTATGATGC GGAACCTCGG 1860  
 GAGCAGGCC TGCGCTTTGG TAGACTCCGC AATGCCACCG CCAGCTGTG GTACAGGCTCT 1920  
 GGGCTGAGG ACCGTGGCTT GGCCTCCCG ATCTTAGATG CAAAGAGTAA GATTGAGCAG 1980  
 ATCCAGCAG TTCTCAGAG CCCCAGCAGT ACAGAGCAGG AGGTGGCTCA GGTGGCCAGT 2040  
 CCTCATCTCT CCTCAGCG AACTCTCCAG GGCCTGAGC TGGATCTGCC CCGTGGAGAG 2100  
 GAGACGTTGT CCTCTCCAG AGACCTGGAG AGTCTTGACA GAAGCTTCAA TGGTCTCTCT 2160  
 ACTATGATC AGAGGAAGAG GAGCAGTTT GAAAAATAA GCAGTCTGA TCCTTCAGGA 2220  
 GCCTTCCGGA TGCTGAGCAC AGCCTACGAG CAGTCAGCCC AGGCTGTCTA GCAGGTCTCC 2280  
 GACAGCTCGC GCCTTTTGGG CCAGCTCAGG GACAGCCGGA GAGAGGCAGA GAGCTGGTG 2340  
 CGGAGGCGCG GAGAGGAGG AGGCACCGGC AGCCCAAGC TTGTGGCCCT GAGGCTGGAG 2400  
 ATGCTCTGCT TGCTGACCT GACACCCACC TTCAACAAGC TCTGTGGCAA CTCCAGGCAG 2460  
 ATGGCTTGA CCCAATATC ATGCCCTGGT GAGCTATGTC CCAAGACAA TGGCACAGC 2520  
 TGTGGCTCCC GCTCAGGGG TGTCTTCCC AGGCCCGGTG GGGCTTCTT GATGGCGGG 2580  
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCAGCTCC AGCGACCAAG GCAGATGATT 2640

AGGCGAGCGG AGGAATCTGC CTCACAGATT CAATCCAGTG CCCAGCGCTT GGAGACCCAG 2700  
GTGAGCGCCA GCCGCTCCCA GATGAGAGAA GATGTGAGAC GCACACGGCT CCTAATCCAG 2760  
CAGGTCCGGG ACTTCCCTAAC AGACCCCGAC ACTGATGCAG CCACTATCCA GGAGGTCAGC 2820  
GAGGCGGTGC TGGCCCTGTG GCTGCCCA CA GACTCAGCTA CTGTTCTGCA GAAGATGAAT 2880  
GAGATCCAGG CCATTGCAGC CAGGCTCCCC AACGTGGACT TGGTGTCTGC CCAGACCAAG 2940  
CAGGACATGG CGCGTGCCCG CCGGTTGCAG GCTGAGGCTG AGGAAGCCAG GAGCCGAGCC 3000  
CATGCAATGG AGGCCAGGTG GGAAGATGTG GTTGGGAACC TGGCGCAGGG GACAGTGGCA 3060  
CTGCAGGAAG CTCAGGACAC CATGCAAGGC ACCAGCGCTT CCCTTCGGCT TATCCAGGAC 3120  
AGGTTGCTG AGGTTTCAACA GGTACTGCGG CCAGCAGAAA AGCTGGTGAC AAGCATGACC 3180  
AAGCAGCTGG GTGACTTCTG GACAAGGATG GAGGAGCTCC GCCACCAAGC CCGGAGCAG 3240  
GGGCGAGAGG CAGTCCAGGC CCAGCAGCTT GCGGAAGGTG CCAGCAGACA GGCATTGAGT 3300  
GCCCAAGAGG GATTGTAGAG AATAAAACAA AAGTATGCTG AGTTGAAGGA CCGTTGGGT 3360  
CAGAGTTCCA TGCTGGGTGA GCAGGCTGCC CGGATCCAGA GTGTGAAGAC AGAGGCAGAG 3420  
GAGCTGTTTG GGGAGACCAT GGAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480  
CTGCGGGGCA GCCAGGCCAT CATGCTGCGC TCGGCGGACC TGACAGGACT GGAGAAGCGT 3540  
GTGGAGCAGA SIQGTGACCA CATCAATGGG CGGTGCTCT ACTATGCCAC CTGCAAGTGA 3600  
TGCTACAGCT TCCAGCCCGT TGCCCCACTC ATCTGCGGCC TTGTCTTTTG GTTGGGGGCA 3660  
GATTGGGTTG GAATGCTTTC CATCTCCAGG AGACTTTCAT GCAGCTTAAA GTACAGCCTG 3720  
GACCAACCTT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAT 3780  
GGGACAGTTA CACTTGACAG ACAAGATGG TGGAGATTGG CATGCCATTG AAACCTAAGAG 3840  
CTCTCAAGTC AAGGAAGCTG GGTGGGCGAG TATCCCCGCG CTTTAGTTCT CCACTGGGGA 3900  
GGAATCCTGG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAATGAA AAGCCAAATA 3960  
AAAATCTTG G

Seq ID NO: 521 Protein sequence  
Protein Accession #: NP\_000219.1

1 11 21 31 41 51  
MRPFFLLCPA LPGLLHAQQA CSRGACYPPV GDLLVGRTRP LRASSTCGLT KPETYCTQYG 60  
EWQMKCKCKD SRQPHNYSH RVENVASSSG PMRWQSQND VNPVSLQLDL DRRPQLQEVN 120  
MEPQGMFPAQ MLIERSSDFG KTWRYVYLA ADCTSTFPV RQGRPQSQWD VRCQSLPQRP 180  
NARLNGGKVQ LNLMDLVSGI PATOSQKIQE VGEITNLRVN FTRLAPVPOR GYHPPSAYYA 240  
VSQRLQGGSC FCHGHADRA PKPGASAGPS TAVQVHDVCV QHENTAGPNC ERCAPFYNNR 300  
PWRPAGGQDA HEQRCRCNG HSETCHFDPA VFAASQAGY GVCNCRDHT EGKNCERCQL 360  
HYFRNRFPQA SIQGTICISE CDFDGAIVGA PCDFVTGQCV CKKHVQGERC DLCKPGPTGL 420  
TYANPQGGCH CDNILGSRD DMPCEDESGR CLCLPNNVGP KCDQCAPYHW KLASGQGCPE 480  
CADCENHSPQ PTQPFVHRAV PCREGFGGLM CSAAAIROCP DRTYGDVATG CRACDCDFRG 540  
TEGPGCDKAS GRCLCRPGLT GPRCDQCQRG YCNRYPVCVA CHEPCQTYDA DLREQALRFG 600  
RLRNATASLW SGPGLEDRGL ASRILDAKSK IEQIRAVLSS PAVTEQEVAV VASAILSLRR 660  
TLQGLQLDL LEEETLSLPR DLESIDRSFN GLLTMYQRKR EQPEKISSAD PSGAFRMLST 720  
AYEQSAQAQA QVSDSSRLLD QLRDSRREAE RLVRQAGGGG GTGSPKLVAL RLEMSLSFDL 780  
TPTFNKLQCN SRQMACTPIS CPGLCLPQDN GTACGSRCRG VLPFRAGAPL MAGQVABQLR 840  
GFNAQLQRT R QMIRAEEESA SIQSSAQLR ETQVSASRSQ MBEDVRTRL LIQVQVDFLT 900  
DDPTDAATIQ EVSEAVLALM LPTDSATVLQ KMHEIQAI AA RLPNVDLVLS QTKQDIARAR 960  
RLQABAEAR SRAGAVEGQV EDVVGNLRRG TVALQEAQDT MGTSRSLRL IQDRVAEVQQ 1020  
VLRPAEKLV T SMTKQLGDFW TRMEELRHQA RQGAABVQA QQLABGASEQ ALSAQEGFER 1080  
IKQKYAELKD RLQSSMLGE QGARIQSVKT EAEELFGETM EMMDRMKIME LEILLRGSQAI 1140  
MLRSADLTGL EKRVEQIRDH INGRVLYAT CK

Seq ID NO: 522 DNA sequence  
Nucleic Acid Accession #: NM\_001944.1  
Coding sequence: 84..3083

1 11 21 31 41 51  
TTTTCTTAGA CATTAACTGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGACTT 60  
TTTACCAGG GAAATCAGAG ACAATGATGG GGCTCTTCCC CAGAACTACA GGGGCTCTGG 120  
CCATCTTGGT GGTGGTGCATA TTGGTTCATG GAGAATTGCG AATAGAGACT AAAGGTCAAT 180  
ATGATGAAGA AGAGATGACT ATGCAACAAG CTAAGAAGAAG GCAAAAACGT GAATGGGTGA 240  
AATTTGCCAA ACCCTGCAGA GAAGGAGAA GATAACTCAA AAGAAACCCA ATTGCCAAGA 300  
TTACTTCAGA TTACCAAGCA ACCCAGAAAA TCACCTACCG AATCTCTGGA GTGGGAATCG 360  
ATCAGCCGCC TTTTGGAAATC TTTGTTGTTG ACAAAAACAC TGGAGATATT AACATAACAG 420  
CTATAGTCGA CCGGGAGGAA ACTCCAAGCT TCCTGATCAC ATGTGCGGCT CTAATGCCCC 480  
AAGGACTAGA TGTAGAGAAA CCACITATAC TAAACGTTAA AATTTTGGAT ATTAATGATA 540  
ATCCTCCAGT ATTTTCACAA CAATTTTCA TGGGTGAAT TGAAGAAAT ACTGCCTCAA 600  
ACTCACTGGT GATGATACATA AATGCCACAG ATGCAGATGA ACCAAACCAC TTGAATTTCTA 660  
AAATTGCGCT CAATTTTCTG TCTCAGGAAC CAGCAGGCAC ACCCATGTTT CTCTTAAGCA 720  
GAAACACTGG GGAAGTCCGT ACTTTGACCA ATTCTCTTGA CCGAGAGCAA GCTAGCAGCT 780  
ATCGTCTGGT TGTGAGTGGT GCAGACAAAG ATGGAAGAAG ACTATCAACT CAATGTGAAT 840  
GTAATATTAA AGTGAAGAGT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900  
CAGCAGTAT TGAAGAAAA ATTTTAAGTT CTGAATTAAT TCGATTTCAA GTAACAGATT 960  
TGGATGAAGA GTACACAGAT AATTGGCTTG CAGTATATTT CTTTACCTCT GGGAAATGAAG 1020  
GAAATTTGTT TGAATACAA ACTGATCCTA GAACTAATGA AGGCATCCTG AAAGTGGTGA 1080  
AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAAACTTAG TATTGCTGTC AAAAACAAAG 1140  
CTGAATTTCA CCAATCAGTT ATCTCTCGAT ACCGAGTTCA GTCAACCCCA GTCAATTTT 1200  
AGGTAATAAA TGTAAGAGAA GGAATTGCAT TCCGTCTGTC TTCCAAGACA TTTACTGTGC 1260  
AAAAAGGCAT AAGTAGCAAA AAATTGGTGG ATTATATCCT GGGAAACATAT CAAGCCATCG 1320  
ATGAGGACAC TAACAAAGCT GCCTCAAATG TCAATATGTT CATGGGACGT AACGATGGTG 1380  
GATACCTAAT GATGATTTCA AAAACTGCTG AAATCAAATT TGTCAAAAAT ATGAACCGAG 1440  
ATTCTACTTT CATAGTTAAC AAAACAATCA CAGCTGAGGT TCTGGCCATA GATGAATACA 1500  
CGGTAATAAC TTTACAGGC ACGTATATG TTAGAGTACC CGATTTCAAT GACAAATTGT 1560  
CAACAGCTGT CCTCGAAAA GATGCAAGTT GCAGTTCTTC ACCTTCGGTG GTTGTCTCCG 1620  
CTAGAACACT GAATATATGA TACATCGGCC CCTATACATT TGCACTGGAA GATCAACCTG 1680  
TAAAGTTGCC TGCCGTATGG AGTATCACA CCCTCAATG TACCTCGGCG CTCCTCAGAG 1740  
CCCAGGAACA GATACCTCCT GGAGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800  
ACAAATCGTG TGAGATGCCA CGCAGCTTGA CACTGGAAGT CTGTCAAGTGT GACACACAGG 1860



GCATCTGTGG AACTTCTTAC CCAACCACAA GCCCTGGGAC CAGGTATGGC AGGCCGCACT 1920  
 CAGGGAGGCT GGGGCTGGCC GCCATGGGCC TGCTGCTCCT TGGTCTCCTG CTGCTGCTGT 1980  
 TGGCCCCCCT TCTGCTGTG ACCTGTGACT GTGGGGCAGG TTCTACTGGG GGAGTGACAG 2040  
 GTGGTTTTT CCCAGTTTCT GATGGCTCAG AAGGAACAAT TCATCAGTGG GGAATTGAAG 2100  
 GAGCCCATCC TGAAGACAA GAAATCACAA ATATTTGTGT GCCTCCTGTA ACAGCCAATG 2160  
 GAGCCGATTT CATGGAAAGT TCTGAAGTTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220  
 TGGAAAGGCAC TTCAAGAAATG GAAATGACCA CTAAGCTTGG AGCAGCCACT GAATCTGGAG 2280  
 GTGCTGACAG CTTTGCACAA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GGAGCAGCCA 2340  
 CTGAGTTGG CATCTGTTC TCAGGGCAGT CTGGAACCAT GAGAACAAGG CATTCCACTG 2400  
 GAGGAACCAA TAAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTCTG GACTCCTACT 2460  
 TTCTTCAGAA AGCATTTGCC TGTGCGGAGG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520  
 TGTGTATCTA TGATAATGAA GCGCAGATG CCACCTGGTC TCCTGTGGGC TCCGTGGGTT 2580  
 GTTGCACTTT TATTGCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640  
 TTAAAAAAT TGCAGAGATA AGCCTTGGTG TTGATGGTGA AGGCAAGAA GTTCAGCCAC 2700  
 CCTCTAAAGA CAGCGGTTAT GGGATTGAAT CCGTGGCCA TCCATAGAA GTCCAGCAGA 2760  
 CAGGATTTGT TAAGTGCAG ACTTTGTCTG AAGTCAAGG AGCTTCTGCT TTGTCCGCT 2820  
 CTGGGTCTGT CCAGCCAGCT GTTCCATCC CTGACCTCT GCAGCATGGT AACTATTAG 2880  
 TAACGGAGAC TTAACCTGCT TCTGTTTCCC TCGTGCAACC TTCCACTGCA GGCTTTGATC 2940  
 CACTTCTCAC ACAAAATGTG ATAGTGACAG AAAGGGTGTG TCTCCCACT TCCAGTGTTC 3000  
 CTGGCAACCT AGCTGGCCCA ACAGAGTAC GAGGGTCACA TACTATGCTC TGTACAGAGG 3060  
 ATCTTTGCTC CCGTCTAATA TGACCAAGAT AGCAAGCTC ACTGTATTGG GCTAATAATT 3120  
 ATCTTTGGAC TAAAGTATTC AAAATAGCAT AGCAAGCTC ACTGTATTGG GCTAATAATT 3180  
 TGGCACTTAT TAGCTTCTCT CATAAATGA TCACGATTAT AAATTAATG TTTGGGTTCA 3240  
 TACCCCAAAA GCAATATGTT GTCACTCTTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA 3300  
 TCTTAAAGTT TTTCAAAACC CTAATATCAT ATTCGC

Seq ID NO: 523 Protein sequence  
Protein Accession #: NP\_001935.1

1 11 21 31 41 51  
 MMGLFPRTTG ALAIFVUVIL VHGLRIETK QYDEEEMTM QAKRRQKRE WVKPAKPCRE 60  
 GEDNSKRNPI AKITSDYQAT QKITRISGV GIDQPPFGIF VVDKNTGDI ITAIVDREET 120  
 PSFLITCRAL NAQGLDVEKP LIITVKILDI NDNPPVFSQQ IFMGEIENS ASNSLVMILN 180  
 ATDADEBNHL NSKIAFKIYS QEPAGTMMFL LSRNTGEVRT LTNSLDRBQA SSYRLVVSQA 240  
 DKDCEGLSTQ CECNIKVKDV NDNFPMFRDS QYSARIEENI LSSELLRFQV TDLDEYTDN 300  
 WLAVYFTTSG NEGNWFETQT DPRTNEGILK VVKALDYEQL QSVKLSIAVK NKAEFHQSVI 360  
 SRYRVQSTPV TIQVINVRG IAFRPASKTF TVQKGISSKK LVDYILSTYQ AIDEDTNKAA 420  
 SNVKYVMGRN DGGYLMDSK TABIKFVKNM NRDSTFVKN TITAEVLAI D EYTGKSTGT 480  
 VYVVRPDPND NCPATVLEKD AVCSSSPSV VSARTLNNRY TGPYTFALD QPVKLPVNS 540  
 ITTNATSLAL LRAQEQIPFG VYHISLVLD SQNNRCMPR SLTLEVCCGD NRGICGTSYP 600  
 TTSPTGYRGR PHSRLCPAA IGLLLGLLL LLLAPLLLT CDCGAGSTGG VTGGFIPVPD 660  
 GSBGTIHQNG IEGAHPEDEK ITNIVPPVT ANGADEMESS EVCINTYARG TAVEGTSME 720  
 MTTLLGAATE SGGAGAFATG TVSGAAGSFG AATGVGICSS QSGSTMTRH STGGTNKDYA 780  
 DGAISMFLD SYFSQKAFAC AEEDDQGEAN DCLLIYDNEG ADATGSPVGS VGCCSFADD 840  
 LDDSLDLSLG PKFKKLAES LGVDGEGKEV QPPSKDSYGV IESCGEPIEV QQTGFVKQQT 900  
 LSGSQGASAL SASGSVQPAV SIPDPLQGN YLVETYSAS GSLVQPSTAG FDPILLQNV 960  
 VTERVICPIS SVPGNLAGPT QLRGSHMLC TEDPCSRLI

Seq ID NO: 524 DNA sequence  
Nucleic Acid Accession #: XM\_058069.2  
Coding sequence: 1..1413

1 11 21 31 41 51  
 ATGAAGTTTC TTCTAATACT GCTCCTGCAG GCCACTGCTT CTGGAGCTCT TCCCTGAAC 60  
 AGCTCTACAA GCCTGGAATA AAATAATGTG CTATTTGGTG AAAGATACTT AGAAAAATTT 120  
 TATGGCCCTTG AGATAAACAA ACTTCCAGTG ACAAATGA ATATAGTGG AAACCTTAATG 180  
 AAGGAAAAAT TCCAAGAAAT GCAGCACTTC TTGGTCTGA AAGTAGCCGG CCACTGGAC 240  
 ACATCTACCC TGGAGATGAT GCACGCACTT CGATGTGGAG TCCCGATGT CCATCATTTT 300  
 AGGGAAATGC CAGGGGGGCC CSTATGGAGG AAACATTATA TCACCTACAG AATCAATAAT 360  
 TACACACCTG ACATGAACCG TGAGGATGTT GACTACGCAA TCCGGAAAGC TTTCCAAGTA 420  
 TGGAGTAATG TTACCCCTTT GAAATTCAGC AAGATTAAAC CAGGCATGGC TGACATTTT 480  
 GTGGTTTTTG CCGTGGAGC TCATGGAGAC TTCATGCTT TTGATGGCAA AGGTGGAATC 540  
 CTAGCCCATG CTTTGGGACC TGGATCTGCC ATTGGAGGGG ATGCACATT CGATGAGGAC 600  
 GAATTCGGGA CTACACATTC AGGAGGCACA AACTTGTTC TCACTGCTGT TCACGAGATT 660  
 GGCCATTCCT TAGGTCTTGG CCATCTAGT GATCCAAAG CCGTAATGTT CCCCACCTAC 720  
 AAATATGTTG ACATCAACAC ATTTGCGCTC TCTGCTGATG ACATACGTGG CATTCACTCC 780  
 CTGTATGGAG ACCCAAAGA GAACCAACGC TTGCCAAATC CTGCAATTC AGAACCACT 840  
 CTCTGTGACC CCAATTTGAG TTTTGTATGT GTCACTACCG TGGGAAATAA GATCTTTTTC 900  
 TTCAAAGACA GGTCTCTCTG GCTGAAGGTT TCTGAGAGAC CAAAGACCAG TGTAAATTTA 960  
 ATTTCTTCTT TATGGCCAAC CTGCGCATCT GGCATTGAAG CTGCTTATGA AATTGAAGCC 1020  
 AGAAATCAAG TTTTCTTTT TAAAGATGAC AAATACTGGT TAATTAGCAA TTTAAGACCA 1080  
 GAGCCAAAT ATCCCAAGAG CATACTTCT TTTGGTTTTC CTAACCTTGT GAAAAAAT 1140  
 GATGCAGCTG TTTTAAACCC ACGTTTTTAT AGGACCTACT TCTTTGTAGA TAACCACTAT 1200  
 TGGAGGTATG ATGAAAGGAG ACAGATGATG GACCTGGT ATCCCAACT GATTACCAAG 1260  
 AACTTCCAAG GAATCGGGCC TAAATTTGAT GCAGTCTTCT ACTCTAAAA CAATACCTAC 1320  
 TATTTCTTCC AAGGATCTAA CCAATTTGAA TATGACTTCC TACTCCACG TATCACCAAA 1380  
 ACACGAAAA GCAATAGCTG GTTTGGTGT TGA

Seq ID NO: 525 Protein sequence  
Protein Accession #: P39900

1 11 21 31 41 51  
 MKPLILLILL ATASGALPLN SSTSLKNNV LFGERYLEKF YGLEINKLPV TKMKYSGNLM 60  
 KEKIQEQHMF LGLKVTGQLD TSTLEMHAP RCGVDVHHF REMPGGPVWR KHYITYRINN 120



YTPDMNREDEV DYAIRKAFQV WSNVTPLEKFS KINTGMADIL VVFARGAHGD FHAFDGKGGI 180  
 LAHAFGPGSG IGGDAHFDEP EFWTTTHSGGT NLPLTAVHEI GHSGLGLGHSS DPKAVMFPTY 240  
 KYVDINTPRL SADDIRIGIQS LYGDPKENQR LFNPDNSEPA LCDPNLSFDA VTTVGNKIPF 300  
 FKDRPFWLKV SERPKTSVNL ISSLWPTLPS GIEAAEYIEA RNQVFLPKDD KYWLISNLRP 360  
 EPNYPKSIHS FGFENFVKKI DAAVFNPRFY RTYFFVDNQY WRDERRQMM DPGYPKLITK 420  
 NFQIGPKID AVFYSKNKY YPFQGSNQFE YDFLLQRITK TLKSNWFGC

Seq ID NO: 526 DNA sequence  
 Nucleic Acid Accession #: NM\_024423.1  
 Coding sequence: 64..2590

1 11 21 31 41 51  
 15 GGCAGGTCTC GCTCTCGGCA CCCTCCCGGC GCCCGCGTTC TCCTGGCCCT GCCCGGCATC 60  
 CCGATGGCGC CCGCTGGGCC CGCGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120  
 CTGACCCCTCG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180  
 CTTCTCTAAC TAGAGGCGA CAAAATAAIT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240  
 TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCACTG 300  
 20 TCACAGCCCA GGGCTGTGTC GCTGCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360  
 GACAAAAGGA AACAGACACA GAAGAGGGT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420  
 TCGAAGACAA GGCACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480  
 ATTCTTGTCT CTATGCAAGA GAATTCCTTG GGCCTTTTCC CATTGTTCT TCAACAAGTT 540  
 GAATCTGATG CAGCACAGAA CTATCTGTC TTCTACTCAA TAAGTGGACG TGGAGTTGAT 600  
 25 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGCACTG GAAATCTATT TTGCACTCGG 660  
 CCTGTGGATG GTGAAGAATA TGATGTTTTT GATTGATTG CTATGCGTC AACTGCAGAT 720  
 GGATATTACG CAGATCTGCC CCTCCCACTA CCGATCAGGG TAGAGGATGA AATGACAAAC 780  
 CACCTGTGTT TCACAGAAGC AATTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840  
 ACTACAGTGG GSGTGGTTG TGCCACAGAC AGAGATGAAC CGGACACAA GCATACGCGC 900  
 30 CTGAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCCTCTTTC TGTGATCCTC 960  
 AGCAGAGGCG TAATCACCAC AGTCTCTCAT TATTTGGACA GAGAGGTTGT AGACAAGTAC 1020  
 TCATTGATAA TGAAGTACA AGACATGGAT GGCAGTTTTT TTGGATTGAT AGGCACATCA 1080  
 ACTTGATACA TAACAGTAAC AGATTCAAAT GATAATGCAC CCACTTTCAG ACAAATGCT 1140  
 TATGAAGCAT TGTGAGAGA AATGCATTG AATGTGAAA TCTTACGAAT ACCTATAGAA 1200  
 35 GATAAGGATT TAATTAACAC TGCCAATTGG AGAGTCAATT TTACCATTTT AAAGGGAAT 1260  
 GAAATGGAGC ATTTCAAAT CAGCACAGAC AAAGAACTA ATGAAGGTGT TCTTTCTGTT 1320  
 GTAAAGCCAC TGAATTATGA AGAAAACCGT CAAGTGAACC TGGAAATTGG AGTAACAAAT 1380  
 GAAGGCCCAT TTGCTAGAGA TATCCCAAGA GTGACAGCCT TGAACAGAGC CTGTGTTACA 1440  
 GTTCATGTGA GGGATCTGGA TGAGGGGCGT GAATGCACCT CTCAGGCCCA ATATGTGCGG 1500  
 40 ATTAAGAAA ACTTAGCAGT GGGTCAAG ATCAACGGCT ATAAGGCATA TGACCCCGAA 1560  
 AATAGAAATG GCAATGGTTT AAGGTACAAA AAATTGCATG ATCCTAAAGG TTGGATCACC 1620  
 ATTGATGAAA TTTCAAGGTC AATCATAACT TCCAAATCC TGGATAGGGA GGTGAAACT 1680  
 CCCAAAAATG AGTTGTATAA TATTACAGTC CTGGCAATAG ACAAAGATGA TAGATCATGT 1740  
 ACTGGAACAC TTGCTGTGAA CATTGAAGAT GTAAATGATA ATCCACAGA AATACTTCAA 1800  
 45 GAATATGTAG TCATTTCGCA ACCAAAAATG GGGTATACCG ACATTTTAGC TGTGTATCCT 1860  
 GATGAACCTG TCCATGGAGC TCCATTTTAT TTCAGTTTGC CCAATACTTC TCCAGAAATC 1920  
 AGTAGACTGT GGAGCCTCAC CAAAGTTAAT GATACAGCTG CCGTCTTTTC ATATCAGAAA 1980  
 AATGCTGGAT TTCAGAAATA TACCATTCTC ATTAAGTATA AAGACAGGCG CGGCCAAGCT 2040  
 GCAACAAAAT TATTGAGAGT TAATCTGTGT GAATGTACTC ATCCAACTCA GTGTCTGCGG 2100  
 50 ACTTCAAGGA GTACAGAGT AATACTTGA AATGGGCAA TCCTTGCAAT ATTACTGGGT 2160  
 ATAGCACTGC TCTTTCTGT ATTGCTAACT TTAGTATGTG GAGTTTTTGG TGCACTAAA 2220  
 GGGAAACGTT TTCTGAAGA TTAGCACAG CAAACTTAA TTATATCAA CACAGAAGCA 2280  
 CCTGGAGAGC ATAGAGTGTG CTCTGCCAAT GGAATTTATGA CCCAACTAC CAACACTCT 2340  
 AGCCAAAGTT TTTGTGATCA TATGGGATCA GGAATGAAAA ATGGAGGGCA GGAACCAAT 2400  
 55 GAAATGATGA AAGGAGGAAA CCAGACCTTG GAATCTGCC GGGGGGCTGG GCATCATCAT 2460  
 ACCCTGGACT CCTGCAGGGG AGGACACAGC GAGGTGGACA ACTGCAGATA CACTTACTCG 2520  
 GAGTGGCACA GTTTTACTCA ACCCGTCTC GGTGAAGAT CCATTAGAGG ACACACTGGT 2580  
 TAAAAATTA ACATAAAGA AATTGCATCG ATGTAATCAG AATGAAGACC GCATGCCATC 2640  
 CCAAGATTAT GTCTCAGCT ATAACTATGA GGAAGAGGGA TCTCCAGCTG GTTCTGTGGG 2700  
 60 CTGCTCAGT GAAAGCAGG AAGAAGATGG CCTTGACTTT TAAATAATT TGGAAACCAA 2760  
 ATTTATTACA TTAGCAGAAG CATGCACAAA GAGATAATGT CACAGTGCTA CAATTAGTTC 2820  
 TTTGTCAGAC ATTCTGAGG TTTCCAAAA TAATATTGTA AAGTTCAT TCAACATGTA 2880  
 TGTATATGAT GATTTTTTTC TCAATTTTGA ATTATGCTAC TCACCAATTT ATATTTTAA 2940  
 AGCCAGTTGT TGCTTATCTT TTCCAAAAAG TGAATAATGT TAAACAGAC AACTGGTAAA 3000  
 65 TCTCAAATC CAGCACTGGA ATTAAGGTCT CTAAGCATC TGCTCTTTT TTTTCTTACG 3060  
 GATATTTTAG TAATAAATAT GCTGGATAAA TATTAGTCCA ACAATAGCTA AGTTATGCTA 3120  
 ATATCACATT ATTATGTATT CACTTTAAGT GATAGTTTAA AAAATAAACA AGAAATATTG 3180  
 AGTATCACTA TGTGAAGAAA GTTTTGGAAA AGAAACAATG AAGACTGAAT TAAATTAATA 3240  
 70 ATGTTGCAGC TCATAAAGAA TTGGGACTCA CCCCTACTGC ACTACCAAT TCATTGACT 3300  
 TTGGAGGCAA AATGTGTTGA AGTGCCCTAT GAAGTAGCAA TTTTCTATAG GAATATAGTT 3360  
 GGAATAAAT GTGTGTGTGT ATATTATTAT TAATCAATGC AATATTAAA ATGAATGAG 3420  
 AACAAAGAGG AAAATGATAA AACTTGAAA TGAGGCTGGG GTATAGTTTG TCCTACAATA 3480  
 75 GAAAAAGAG AGAGCTTCTC AGGCTTGGGC TCTTAAATGC TGCAATATA CTGAGTCTAT 3540  
 GAGGAAATAG TTCTGTCCA ATTTGTGTA TTTGTTTAAA ATTGTAAATA AATTAACTT 3600  
 TTCGTGTTTC TGTGGGAAGG AAATAGGGA TCCAATGGAA CAGTAGCTTT GCTTTGCACT 3660  
 CTGTTTCAAG ATTTCTGCAT CCACAAGTTA GTAGCAAACT GGGGAATACT CGCTGCAGCT 3720  
 GGGTTCCCT GCTTTTGTGT AGCAAGGTC CAGAGATGAG GTGTTTTTT CGGGGAGCTA 3780  
 80 ATAACAAAA CATTTTAAA CTTACCTTTA CTGAAGTTAA ATCCTCTATT GCTGTTCTA 3840  
 TTCTCTCTTA TAGTACCAA CATCTTTTTA ATTTAGATCC AAATAACCAT GTCCTCCTAG 3900  
 AGTTTAGAGC CTAGAGGGAG CTGAGGGGAG GATCTTACTG AAGCACCCT GGGGAGATTG 3960  
 ATTGTCCTTA AACTTAAGCC CCACAACCTT GACACCTGAT CAGGCTGGG AGCTACAAAA 4020  
 85 TTTCATTTT CTCTCACTG CCTTCTTCT GAGTGGCATT GGCCTGAATC AAGGAAAGCC 4080  
 AGGCTTGTG GSCCCCTTC TTTGGGCTTT CTGCTAAAGC AACACCTCCA GCAGAGATTG 4140  
 CCTTAAGTGA CTTCAAGTTT TCCACCATCC TTCAGCGTGA ATTAATTTT AATCAGTTG 4200  
 CTTTCTCCAG AGAAATTTTA AAATAATAGA AGAAATAGAA ATTTGAAATG TATAAAGAA 4260  
 AAAGATCAAG TTTGCTATT AGAACAGAGG GAACCTTGGG AGAAAGCAGC CCAAGTAGGT 4320  
 TATTTGTACA GTCAAGGGC AACAGGAAGA TGCAGGCCCT CAAGGGCAAG GAGAGGCCAC 4380  
 AAGGAATATG GGTGGAGTA AAAGCAACAT CGTCTGCTTC ATACTTTTTC CTAGGCTTGG 4440

CACTGCTTT TCCTTTCTCA GGCCAAATGGC AACTGCCATT TGAGTCCGGT GAGGGATCAG 4500  
 CCAACCTCTT CTCTATGGCT CACCTTATIT GGAGTGAGAA ATCAAGGAGA CAGAGCTGAC 4560  
 TGCAATGATGA GTCTGAAGGC ATTTGACAGA TGAGCCTGAA CTGGTTGTGC AGAACAACAA 4620  
 AGGCATTAT CAT GGGAAATTGTT GTATTCTCTC TGCAGCCCTC CTCTGGGCA CTAAGAAGGT 4680  
 CTATGAATTA AATGCCCTATC TAAATTTCTG ATTTATTCCT ACATTTCTG TTTTCTAATT 4740  
 TGACCCATAA ATCTATGTGT TTTAGACTTA GACTTTTAT TGCCCCCCCC CCCTTTTITT 4800  
 TTGAGACGGA GTCTCGCTCT GAGGCACAGG CTGGAGTGCA GTGGCTCCGA TCTCTGCTCA 4860  
 CTGAAAGCTC CGCTCCCGG GTTCATGCCA TTCTCTGCTC TCAGCCTCTC GAGTAGCTGG 4920  
 GACTACAGGC GCCCACACG ACGCCCGGCT AATTTTGTGT ATTTTAAATA GAGACGGGGT 4980  
 TTCACTGTGT TAGCCAGGAT GGCTCGATC TCCTGACCTC GTGATCCGCC TGCTCGGCC 5040  
 TCCCAAAGTG CTGGGATTAC AGGCATGACC CACCGCTCCC GGCTTGTTT TCCGTTTAAA 5100  
 TCGCTCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTGTGATCAT ACGAATTGGA 5160  
 TCAATCTTGA AATACTCAAC CAAAGACAG TCGAGAAGCC AGGGGGAGAA AGAAGCTCAGG 5220  
 GCACAAAATA TTGGTCTGAG AATGGAATTC TCTGTAAGCC TAGTGTCTGA AATTTCTGTC 5280  
 TGTAACAGCA AGGCAGTTT ATCTAACGGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340  
 CCACTACACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTGCCG ATAAACTTTC 5400  
 TCAAGAGCA ACCAGTATCA CTTCCTGTGT TATAAAACCT CTAACCATCT CTTTGTCTCT 5460  
 TGAACATGCT GAAACACACC TGGTCTGCAT GTATGCCCGA ATTTGTAATT CTTTCTCTC 5520  
 AAATGAAAAT TTAATTTTAG GGATTCAATT CTATATTTT ACATATGTAG TATTATTATT 5580  
 TCCTTATATG TGTAAAGGTGA AATTTATGGT ATTTGAGTGT GCAAGAAAAT ATATTTTAA 5640  
 AGCTTTCATT TTTCCCCAG TGAATGATTT AGAATTTT ATGTAATAT ACAGAATGTT 5700  
 TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAATGTCAG TGGGTTTGT TTTGCAATGT 5760  
 TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAAACTTGG 5820  
 CTGCTTAAAA TAAGCAAAAA TTGGATGCAT AAGTAATAT TTACAGATGT GGGGAGATGT 5880  
 AATAAACAAA TATTAACCTG GCTGCTTAAA ATAAGCAAAA ATTGGATGCA TAAAGTAATA 5940  
 TTTACAGATG TGGGGAGATG TAATAAACA ATATTAACCT GGTTTCTGT TTTTCTGTA 6000  
 TTTAGAGATT AAATAATCTT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCTGGCATG 6060  
 GAAATAGAAA TACTCAATTA TGTCTTTGTT GTATTAATGG GGAATATTTT GACCAATGTT 6120  
 TCATTATCAA ATTTGTGACA TCATTAAAT ATATTTGAAT GTTGGGAAGA GACTCACTATT 6180  
 TTGAAGCACA GCTTTACAGA TGAGTATCTA TGATACATAT GTATAATAA TTTTGTATGG 6240  
 GTATTAAAG TTTTGAAGG TGGTTATAAT TGCAGATAT TCCATGAATA GTACACTGAC 6300  
 ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGAAA ACATGAGTTA AAAGAAAAG 6360  
 CAGGCAATAT TGCAGTCTG ATTTCTGCCA TTACAGGATA GATAATGCCCT GAACCTTAAT 6420  
 GACAAGATGA TCCAACATA AAGTGTCTCT GTGCTTACA GTGAATCTTT TCCCATGCA 6480  
 GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTCA AAAATCTATT AGCTATATCA 6540  
 AAAGCCTTAC ATTTTAAAT AGGTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600  
 ACCATTATTT TTGTGTATGT CTTCAAGAA GTTTCATTGA TTTTGTGTT TAATAGTAAA 6660  
 ATACCGGATA CATTTCACTG TCCCTTCAGT ATTGATTGG TTGAATATTG GTGCATAATG 6720  
 GTTGAGAAGC ATGGACACTA GAGCCAGAA GCTTGGATAT GAATCCTGGA TCTGTCACTT 6780  
 ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTCTC ATTAATAATCA 6840  
 ATGAACAATG CCGAGCTCAT GGGGTGTTG AATGATTAAA TTAGTTAATA TACCTAAAGT 6900  
 ACATAGAAC CAAGCTGCAC ATAGTAAAAG AATTATAAGT GTGAGGTAGT TGGTAAAT 6960  
 ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAGTTTGTG 7020  
 CATATATATA ATCCCGAAAC ATG

Seq ID NO: 527 Protein sequence  
 Protein Accession #: NP\_077741.1

1 11 21 31 41 51  
 MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEBCFRS 60  
 ADLIRSSDDP FRVLNDGSVY TARAVALSDK KRSPTIWLSD KRKQTQKEVT VLEHQKQVS 120  
 KTRTRETULV RRARRRWAPI PCSMQENSLG PFLFLQOQVE SDAAQNYTVF YSISGRGVDK 180  
 EPLMLFYIER DTGNLFCTRP VDREEDVDVD LIAYASTADG YSADLPLPLP IRVEDENDNH 240  
 PVFTEAIYNF EVLESSRPGT TVGVVCAADR DEPDTHMLTR KYSLQOTPR SPGLFSVHPS 300  
 TGVITTVSHY LDREVVDKYS LIMKVQDMDG QPFLIGTST CIIITVDSND NAPTFRQNAV 360  
 EAFVEENAFN VEILRIPIED KDILINTANR VNPTILKQNE NGHFKISTDR ETNEGVLSV 420  
 KPLNYEENRQ VNLIGVNNIE APFARDIPRV TALNRALVTV HVRDLDEGPE CTCAAQYVRI 480  
 KENLAVGSKI NGVKAYDPEN RENGNGLEYKK LHDPKGMITI DEISGSIITS KILDRVETP 540  
 KNELYNITVL AIDKDRSCT GTLAVNIEDV NDNPPILLQE YVVICRPMKG YTDILAVDSD 600  
 EPVHGAPFFY SLPNTSPEIS RLWSLTKVND TAARLSYQKN AGQFEYTIPI TVKDRAGQAA 660  
 TKLELVNLCE CTHPTQCRAT SRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720  
 KRFPEDLAQQ NLIISNTEAP GDDRVCSSANG FMTQTTNNSS QGFCGTMGSG MKNGGQETIE 780  
 MMKGGNQTLB SCRAGHHT LDSCRGGHTE VDNCRITYSE WHSFTQPRLG BESIRGHTG

Seq ID NO: 528 DNA sequence  
 Nucleic Acid Accession #: NM\_001941.2  
 Coding sequence: 64..2754

1 11 21 31 41 51  
 GGCAGGTCTC GCTCTCGGCA CCTTCCCGGC GCCCGCGTTC TCCCTGGCCCT GCCCGGCATC 60  
 CCGATGGCCG CCGCTGGGCC CCGGCGCTCC GTGCGCGGAG CCGTCTGCCT GCATCTGCTG 120  
 CTGACCCCTC TGATCTTCAG TCGTGATGTT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180  
 CCTTCTAAAC TAGAGGCAGA CAAAATAATT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240  
 TCTGCAGACC TCATCCGGTC AAGTGATCCT GATTTCAGAG TTCTAAATGA TGGGTCACTG 300  
 TACACAGCCA GGGCTGTGTC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360  
 GACAAAAGGA AACACACACA GAAAGAGGTT ACTGTGCTGC TAGAACATCA GAAGAAGGTA 420  
 TCGAAGACAA GACACACTAG AGAAAAGTGT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480  
 ATTCTTGTCT CTATGCAAGA GAATTCCTTG GGCCTTTTCC CATTTGTTCT TCAACAAGTT 540  
 GAATCTGATG CAGCACAGAA CTATCTGTCT TTCTACTCAA TAAGTGGAGC TGGAGTTGAT 600  
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTATT TTGCACTCGG 660  
 CCGTGGGATC GTGAGAATA TGAATTTTTT GATTTGATTC CTTATGCTGC AACTGCAGAT 720  
 GGATATTTCAG CAGATCTGCC CCTCCCACTA CCCATCAGGG TAGAGGATGA AAATGACAAC 780  
 CACCTCTGTT TCACAGAGGC AATTTATAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840  
 ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGCCG 900  
 CTGAAATACA GCATTTTGCA GCAGACACCA AGGTCACCTG GGCTCTTTTC TGTGCATCCC 960

	AGCACAGGCG	TAATCACCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
	TCATTGATAA	TGAAAGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTGAT	AGGCACATCA	1080
	ACTTGTATCA	TAACAGTAA	AGATTCAAAT	GATAATGCAC	CCACTTTCAG	ACAAAATGCT	1140
5	TATGAAGCAT	TTGTAGAGGA	AAATGCATT	AAATGTGAAA	TCTTACGAAT	ACCTATAGAA	1200
	GATAAGGATT	TAATTAACAC	TGCCAATTGG	AGAGTCAATT	TTACCAATTT	AAAGGGAAT	1260
	GAAAAATGAC	ATTTCAAAT	CAGCACAGAC	AAAGAACTA	ATGAAGGTTG	TCTTCTGTT	1320
	GTAAGGCCAC	TGAATTATGA	AGAAAACCGT	CAAGTGAACC	TGGAAATTGG	AGTAAACAAT	1380
	GAGCGGCCAT	TTGCTAGAGA	TATTCGCCAG	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
10	GTTCATGTGA	GGGATCTGGA	TGAGGGGCGT	GAATGCACCT	CTGCAGCCCA	ATATGTGGGG	1500
	ATTAAGAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAAATCC	TGGATAGGGA	GTTTGAACCT	1680
	CCCAAAATAG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
15	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
	GAATATGTAG	TCATTTGCAA	ACCAAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGT	TCCATTTTAT	TTGAGTTTGC	CCAATACCTC	TCCAGAAATC	1920
	AGTAGACTGT	GGAGCCTCAC	CAAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGATA	TACCATTCTC	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
20	GCAACAAAAT	TTGTAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACTCA	GTGTCTGGCG	2100
	ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCTTTCTGCT	ATTGCTAAT	TTAGTATGTG	GAGTTTTGGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCTGAAGA	TTTAGCACAG	CAAACTTAA	TTATATCAA	CACAGAAACA	2280
	CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAACTTAC	CAACACTCT	2340
25	AGCCAAGGTT	TTTGTGGTAC	TATGGGATCA	GGAATGAAAA	ATGGAGGGCA	GGAACCAATT	2400
	GAATATGATA	AAGGAGGAAA	CCAGACCTTG	GAATCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCTCGACT	CTGCAAGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACCTACTCG	2520
	GAGTGGCACA	GTTTTACTCA	ACCCGCTCTC	GGTGAATAAT	TGCATCGATG	TAATCAGAA	2580
	GAGAGCCGCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGACTC	2640
30	CCAGCTGGTT	CTGTGGGCTG	CTGCACTGAA	AAGCAGGAAG	AAGATGGCCT	TGACTTTTAA	2700
	AAATAATTTG	AACCCAAAT	TATTACATTA	GCAGAAGCAT	GCACAAAGAG	ATAATGTAC	2760
	AGTGCTACAA	TTAGGCTCTT	GTCAACATTT	CTGGAGGTTT	CCAAAATAAA	TATTGTAAAG	2820
	TTCAATTTCA	ACATGTATGT	ATATGATGAT	TTTTTCTCTA	ATTTTGAATT	ATGCTACTCA	2880
	CCAAATTTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAATGTAA	2940
35	AACAGACAA	TGGTAAATCT	CAAACTCCAG	CAGTGAAT	AAGGTCTCTA	AAGCATCTGC	3000
	TCCTTTTTTT	TTTTACGAGT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TAGCTATATA	TCACATTAAT	ATGTATTCAC	TTTAAGTGAT	AGTTTAAAAA	3120
	ATAAACAGAA	AATATTAGAT	ATCACTATGT	GAAGAAAGTT	TTGGAAGAGA	AACAATGAAG	3180
	ACTGAATTA	ATTAAAAATG	TTGCAGCTCA	TAAAGAAATG	GGACTCAACC	CTACTGACT	3240
40	ACCAAAATCA	TTTGACTTTG	GAGGCAAAAT	GTGTTGAAGT	GCCCTATGAA	GTAGCAATTT	3300
	TCATATAGGA	TATAGTTGGA	AAATAATGTT	TGTGTGTATA	TTATTATTA	TCAATGCAAT	3360
	ATTTAAATG	AAATGAGAAA	AAAGAGGAAA	ATGTTAAAAA	CTTGAATGAA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGCTGC	3480
	ATTATAACTG	AGTCTATGAG	GAATAGTTTC	CTGTCCAATT	TGTGTAATTT	GTTTAAAAAT	3540
45	GTAAATAAAT	TAAACTTTTC	TGGTTTCTGT	GGGAAGGAAA	TAGGGAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATACCTGC	TGCAGCTGGG	GTTCCCTGCT	TTTTGTAGC	AAGGGTCCAG	AGATGAGGTG	3720
	TTTCTTGGG	GGAGCTAATA	ACAAAACAT	TTTAAAACTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTCTTATATG	TCTCTTATAG	TGACCAACAT	CTTTTAAATT	TAGATCCAAA	3840
50	TAACCATGTC	CTCCTAGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCCCTGG	GAGATTGATT	GTCTTAAAC	CTAAGCCCCA	CAAACTTGAC	ACCTGATCAG	3960
	GTCTGGGAGC	TACAAAATTT	CATTTTCTC	CTCACTGCC	TTCTTCTGAG	TGGCATTTGG	4020
	CTGAATCAAG	GAAAGCCAGG	CCTTGTGGGC	CCCTTCTTT	CGGCTTCTG	CTAAGGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAAGTGACTC	CAGGTTTCC	ACCATCCTTC	AGCGTGAATT	4140
55	AATTTTAAAT	CAGTTTGCTT	TCTCCAGAGA	AATTTTAAAA	TAATAGAAGA	AATAGAAATT	4200
	TTGAATGTAT	AAAAGAAAAA	GATCAAGTTG	TCATTTTAGA	ACAGAGGGA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTAGGTTAT	TTGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCCTCAA	4320
	GGGCAAGGAG	AGGCCCAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCATA	4380
	CTTTTCTCTA	GGCTTGGCAC	TGCTTTTCC	TTTCTCAGGC	CAATGGCAAC	TGCCATTTGA	4440
60	GTCCGGTGAG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTGGA	GTGAGAAATC	4500
	AAGGAGACAG	AGCTGACTGC	ATGATGAGTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTGTGTCAGA	ACAAACAAGG	CATTATGGG	AAATGTTGTA	TTCTTCTGTC	AGCCCTCCTT	4620
	CTGGGCACTA	AGAAGGTCTA	TGAATTAAT	GCCTATCTAA	AATTTCTGAT	TATTTCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAAAATC	TATGTGTTT	AGACTTAGAC	TTTTTATTGC	4740
	CCCCCCCC	TTTTTTTT	AGACGGAGTC	TGCTCTGAC	GCACAGGCTG	GAGTGCACTG	4800
65	GCTCCGATCT	CTGCTCAGTG	AAAGCTCCGC	CTCCGGGTT	CATGCCATTC	TCCTGCCTCA	4860
	GCTCTCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCACG	CCCGGCTAAT	TTTTTGTATT	4920
	TTTAATAGAG	ACGGGGTTTC	ACTGTGTAG	CCAGGATGTT	CTGATCTCC	TGACCTCGTG	4980
	ATCCGCTCTG	CTCGGCCCTC	CAAAGTGCTG	GGATTACAGG	CATGACCCAC	CGCTCCCGGC	5040
	CTGTTTCTCC	GTTTAAAGTC	GTCTTCTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100
70	TGATCATACG	AATTTGATCA	ATCTTGAAAT	ACTCAACCAA	AAGACAGTCG	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGCA	CAAAATATTG	GTCTGAGAA	GGAATTTCTC	GTAAGCCTAG	5220
	TTGCTGAAAT	TTCTGTGCT	AACCAGAGGC	CAGTTTATC	TAACGGCTAC	TGAACACCC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
	CTAGTGCCGA	TAACTTTCT	CAAGAGCA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
75	TAACCATCTC	TTTGTCTTT	GAACATGCTG	AAAACCACT	GGTCTGATG	TATGCCCGAA	5460
	TTTGTAAATC	TTTTCTCTCA	AATGAAATTT	TAATTTTAGG	GATTCAATTC	TATATTTTCA	5520
	CATATGTAGT	ATTATTATTT	CCTTATATGT	GTAAGGTGAA	ATTTATGGTA	TTTGAAGTGTG	5580
	CAAGAAAAATA	TATTTTAAA	GCTTTCATTT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
	TGTAATAATA	CAGAATGTTT	TTTCTTACTT	TTATAAGGAA	GCAGCTGTCT	AAAATGCACT	5700
80	GGGGTTGTT	TTGCAATGTT	TTAAACAGAG	TTTTAGTATT	GCTATTAAAA	GAAGTTACTT	5760
	TGCTTTTAAA	GAAACTTGGC	TGCTTAAAT	AAGCAAAAT	TGGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAAT	ATTAACCTGG	TTTCTTGT	TTGCTGTATT	5880
	TAGAGATTAA	ATAATTCTAA	GATGATCACT	TGCAAAAT	ATGCTTATGG	CTGGCATGGA	5940
	AATAGAAATA	CTCAATTATG	TCTTTGTGTT	ATTAATGGGG	AATATTTTGG	ACAATGTTTC	6000
85	ATTATCAAT	TGTGCACATC	ATTAATATAT	ATTGTAATGT	TGGGAAGAGA	TCACTATTTT	6060
	GAAGCACAGC	TTTACAGATG	AGTATCTATG	ATACATATGT	ATAATAAAT	TTGATCGGGT	6120
	ATTAAGAAATA	TTAGAAGGTG	GTTATAATTT	CAGAGTATTC	CATGAATAGT	ACACTGACAC	6180

WO 02/086443

PCT/US02/12476

AGGGGTTTTA CTTTGAGGAC CAGTGTAGTC AAGGGAAAAC ATGAGTTAAA AAGAAAAGCA 6240  
 GGCNAATATG CAGTCTTGAT TCTGCCACTT ACAGGATAGA TAATGCCTGA ACTTTAATGA 6300  
 CAAGATGATC CAACCAATAA GGTGCTCTGT GCTTCACAGT GAATCTTTTC CCCATGCAGG 6360  
 AGTGTGCTCC CTTACAAACG TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAAA 6420  
 AGCCTTACAT TTTAATATAG GTTGAACCAA AATTTCAATT CCAAGTAACTT CTATTGTAAC 6480  
 CATATTATTT GTGTATGTCT TCAAGAAATG TCATTGGATT TTTGTTTGTG ATAGTAAAAAT 6540  
 ACCGATACAT TTTCACGTGT CCTTCAGTAT TGATTGGGT GAATATTGGG TCATATATGGT 6600  
 TGAGAAGCAT GGCACATAGA GCCAGAATGC TTGGATATGA ATCCTGGATC TGTCACCTAC 6660  
 TTCTGTGTGA CCTTTGAAAG GCTACTTATT TCCTCTCTTA GCTTTCTCAT TAAAATCAAT 6720  
 GAACRAATGCC AGCCTCATGG GGTGTGTGAA TGATTAAATT AGTTAATATA CCTAAGATAC 6780  
 ATAGAACACT GCCTGCACAT AGTAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840  
 GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTT AGGGAATAA AGTTTGTGCA 6900  
 TATATATAAT CCGAAACAT G

Seq ID NO: 529 Protein sequence  
 Protein Accession #: NP\_001932.1

1 11 21 31 41 51  
 | | | | |  
 20 MAAAGPRRSV RGAVCLHLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECPRS 60  
 ADLIRSSDPD PRVLNDGSVY TARAVALSDK KRSPTIWLSD KRKQTQKEVT VLEHQKKVS 120  
 KTRHRTRETVL RRAKRWAPI PCSMQENSLG PFPLFLQQVE SDAAGNTYVF YSISGRGVDK 180  
 EPNLIFYIER DTGNLFCTRP VDREEDVDFD LIAYASTADG YSADLPLPLP IRVEDENDNH 240  
 25 PVTEAIYNF EVLESSRPGT TVGVVCAADR DEPDTHMTRL KYSILQOTFR SPGLPSVHPS 300  
 TGVITTVSHY LDREVVDKYS LIMKVQDMDG QFFGLIGTST CIITVDSND NAPTFRQNAV 360  
 EAFVEENAFN VEILRIPIED KDLINTANWR VNFITLKONE NGHFRISTDK ETNEGVLVSV 420  
 KPLNVEENRQ VNLEIGVNNV APPARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480  
 KENLAVGSKI NGYKAYDPEN RENGNGLYKK LHDPCGWITI DEISSGIITS KILDREVETP 540  
 30 KNELYNITVL AIDKDRSCT GTLAVNIEDV NDNPPETLQE YVVICPKPMG YTDILAVDPD 600  
 EPVHGAPFFP SLPNTSPSIS RLWSLTKVND TAARLSYQKV AGFQZYTPII TVKDRAGQAA 660  
 TKILRVNLCE THPTQCRAT SRSSTGVLGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720  
 KRFPEDLAQQ NLIISNTEAP GDDRVCANG FMTQTNNSS QGFCGTMGSG MNKGQETIE 780  
 MMKGNQTL ESCRGAHHHT LDSCRGGHTE VDNCRYTYSE WHSFTQPLRG EKLHRCNQNE 840  
 35 DRMPQDVVL TYNYEGRGSP AGSVGCCSEK QEBDGLDFLN NLEPKFITLA EACTKR

Seq ID NO: 530 DNA sequence  
 Nucleic Acid Accession #: NM\_016583.2  
 Coding sequence: 72..842

1 11 21 31 41 51  
 | | | | |  
 40 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60  
 TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCGAGA 120  
 45 CCATGGCCCA GTTTGAGGCG CTGCCGCTGC CCTGGACCA GACCCTGCC TGAATGTGA 180  
 ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GOCCTCAGCA 240  
 ATGGCCTGCT GTCTGGGGGC CTGTGGGCA TTCTGGAATA CCTTCGCTC CTGGACATCC 300  
 TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGA AAGTGACGT 360  
 50 CAGTGATTCC TGGCCTGAAC AACATCAITG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420  
 AACTTGGCCT TGTGCAGAGC CCTGATGGCC ACGTCTCTTA TGTACCATC CCTCTCGGCA 480  
 TAAAGTCCA AGTGAATACG CCCCTGGTCG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC 540  
 TGGACATCAC TGCAGAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGCTCC 600  
 TTGGTGACTG CACCCATTCC CCTGGAAGCC TGCAAAATTC TCTGCTTGAT GGACTTGGCC 660  
 55 CCCTCCCATC TCAAGTCTT CTGACAGGCC TCACAGGGAT CTTGAATAAA GTCTGCTG 720  
 AGTTGGTTCA GCGCAACGTC TGCCCTCTGG TCAATGAGGT TCTCAGAGGC TTGGACATCA 780  
 CCCTGTGTGA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTG ATCAAGGTCT 840  
 AAGCCTTCCA GGAAGGGGCT GGCCTCTGCT GAGCTGCTTC CCACTGCTCA CAGATGCTG 900  
 GCCCATGTGC TGAAGATGTA CACAGTTGCC TTCTCTCGA GGAACCTGCC CCCTCTCCTT 960  
 60 TCCACACAGG CGTGTGTAAC ATCCCATGTG CCTCACCTAA TAAATGGCT CTCTCTCTGC 1020  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 531 Protein sequence  
 Protein Accession #: NP\_057667.1

1 11 21 31 41 51  
 | | | | |  
 65 MPOTGGLIVF YGLLAQMAQ FGGLPVPLDQ TLPLNVNPAI PLSPTGLAGS LTNALSNGLL 60  
 SGGLLGILEN LPLLDILKPG GGTSGGLLGG LLGKVTSPVP GIANNIIDIKV TDPQLLELGL 120  
 70 VQSPDGHRLY LTIPLGIKIQ VNTPLVGASL LRLAVLKDIT AEILAVRDQK ERIHLVLGDC 180  
 THSPGSLQIS LLDGLGLPLI QGLDLSLTGI LNKVLPPLVQ GNVCPVNEV LRGLDITLVH 240  
 DIVNMLIRGL QFVIKV

Seq ID NO: 532 DNA sequence  
 Nucleic Acid Accession #: NM\_004363.1  
 Coding sequence: 115..2223

1 11 21 31 41 51  
 | | | | |  
 80 CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60  
 TCTTGAAGCT CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120  
 TCTCCCTCGG CCCCTCCCA CAGATGGTGC ATCCCTGGC AGAGGCTCCT GCTCACAGCC 180  
 TCACTTCTAA CCTTCTGGAA CCGGCCACCC ACTGCCAAGC TCACTATTGA ATCCAGCCCG 240  
 85 TTCAATGTGG CAGAGGGGAA GGAGGTGCTT CTACTTGTCC ACAATCTGCC CCAGCATCTT 300  
 TTTGGCTACA GCTGGTACAA AGGTGAAGA GTGGATGGCA ACCGTCAAT TATAGGATAT 360  
 GTAATAGGAA CTCACAAGC TACCCAGGG CCCGCATACA GTGGTGAGA GATAATATAC 420  
 CCCAATGCAT CCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCTTA 480

5  
10  
15  
20  
25  
30  
35  
40  
45

```

CACGTCATAA AGTCAGATCT TGTGAATGAA GAAGCAACTG GCCAGTTCGG GGTATACCGG 540
GAGCTGCCCA AGCCCTCCAT CTCCAGCAAC AACTCCAAAC CCGTGGAGGA CAAGGATGCT 600
GTGGCCTTCA CCGTGAAGCC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAA 660
CAGAGCTCC CCGTCAGTCC CAGGCTGCAG CTGTCCAATG GCAACAGGAC CCTCACTCTA 720
TTCAATGTCA CAAGAAATGA CACAGCAAGC TACAATGTG AAACCCAGAA CCCAGTGAGT 780
GCCAGGCGCA GTGATTCTAGT CATCTCTAAT GTCCTCTATG GCCCGGATGC CCCCACTT 840
TCCCTCTTAA ACACATCTTA CAGATCAGGG GAAAATCTGA ACCTCTCTG CCACCGAGCC 900
TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTTCCAGCA ATCCACCCAA 960
GAGCTCTTTA TCOCACAAAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020
AACTCAGACA CTGGCTCAA TAGGACCAAC GTCAAGCAGA TCACAGTCTA TGCAGAGCCA 1080
CCCAAAACCT TCATCACCAG CAACAACCTC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140
TTAACTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200
GTCCCGTCA GTCCAGGCT GCAGCTGTCC AATGACAACA GGACCTCAC TCTACTCAGT 1260
CTCACAAGGA ATGATGTAGT ACCCTATGAG TGTGGAATCC AGAACGAATT AAGTGTGAC 1320
CAGAGCAGCC CAGTCACTCT GAATGTCTCT TATGGCCAG ACAGCCCACT CATTTCCCCC 1380
TCATACACCT ATTACCGTCC AGGGGTGAAC CTCAGCTCT CCTGCCATGC AGCCTCTAAC 1440
CCACCTGCAC AGTATTCTTG GCTGATTGAT GGAACATCC AGCAACACAC ACAAGAGCTC 1500
TTTATCTCCA ACATCACTGA GAAGAACAGC GGACTCTATA CTGCCAGGC CAATAACTCA 1560
GCCAGTGGCC ACAGCAGGAC TACAGTCAAG ACAATCAGAG TCTCTGCGGA GCTGCCAAG 1620
CCCTCCATCT CAGCAACAA CTCCAAACCC GTGGAGGACA AGGATGCTGT GGCCTTCACC 1680
TGTGAACCTT AGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTA GAGCCTCCCA 1740
GTCAGTCCCA GGTGACGCT GTCCAATGGC AACAGACCC TCACTCTATT CAATGTACA 1800
AGAAATGAGC CAAGAGCCTA TGTATGTGGA ATCCAGAACT CAGTGAAGTC AAACCGCAGT 1860
GACCCAGTCA CCTGATGTG CCTCTATGGG CCGGACACCC CCATCATTTT CCCCCAGAC 1920
TCGTCTTACC TTTGGGAGC GAACCTCAAC CTCTCTGCC ACTCGGCTC TAACCCATCC 1980
CCGCAATATT CTTCGCGTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTATC 2040
GCCAAATCA CGCAAAATTA TAACGGGACC TATGCCTGTT TGTCTCTTAA CTGTGGTACT 2100
GGCCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TTCTCTGGT 2160
CTCTCAGCTG GGGCACTGT CGGCATCATG ATTGGAGTGC TGGTGGGGT TGCTCTGATA 2220
TAGCAGCCCT GGTAGTATT CTTCATTTC GGAAGACTGA CAGTGTGTTT GCTTCTCTCT 2280
TAAAGCAATT GCACAGCTA CAGTCTAAA TTGCTTCTT ACCAAGGATA TTTACAGAAA 2340
AGACTCTGAC CAGAGATCGA GACCATCTA GCCAATCAG TGAACCCCA TCTCTACTAA 2400
AAATACAAA ATGAGCTGGG CTGTGGTGGC CGCACCTGTA GTCCAGTTA CTCGGGAGG 2460
TGAGGCAGGA ATAGCTCTG AACCCGGGAG GTGAGATTG CAGTGAGCCC AGATCGCACC 2520
ACTGCACTCC AGTCTGGCAA CAGAGCAAGA CTCCATCTCA AAAAGAAAAG AAAAGAAGAC 2580
TCTGACCTGT ACTCTGAAAT ACAAGTTTCT GATACCACTG CACTGTCTGA GAATTTCCAA 2640
AACTTTAATG AACTAACTGA CAGCTTCATG AAAGTGTCCA CCAAGATCAA GCAGAGAAAA 2700
TAATTAATTT CATGGGACTA AATGAACATA TGAGGATTGC TGATTCTTTA AATGTCTTGT 2760
TTCCAGATT TCAGGAAACT TTTTCTTTT TAAGCTATCC ACTCTACAG CAATTGTGATA 2820
AAATATACTT TTGTGAACAA AAATTGAGAC ATTTACATT TCTCCCTATG TGGTGGCTCC 2880
AGACTTGGGA AACTATTAT GAATATTAT ATTGTATGTT AATATAGTTA TTGCACAAGT 2940
TCAATAAAAA TCTGCTCTT GTATAACAGA AAAA

```

Seq ID NO: 533 Protein sequence  
Protein Accession #: NP\_004354.1

50  
55  
60

```

1 11 21 31 41 51
| | | | |
MESPSAPPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVABGKE VLLLVHNLPO 60
HLFGYSWYKG ERVDGNRQII GYVIGTQOAT PGPAYSGREI IYFNASLLIQ NIIQNDTGFY 120
TLEVIKSDLV NBEATGQFRV YPELPKPSIS SNNSKPVEDK DAVAPTCEPE TQDATYLMWV 180
NNQSLFVSFR LQLSNGNRTL TLFNVTRNDT ASYKCTQNP VSARRSDSVI LNVLYGPDAP 240
TISPLNTSYR SGENLNLSCH AASNPPAQYS WFNVTGTFQS TQELFIPNIT VNNSGSYTCQ 300
AHNSDTGLNR TTVTITYYA EPKPFITSN NSNPVEDEDA VALTCEPEIQ NTLYLWVWNN 360
QSLFVSPRLQ LSNDRNTLTL LSVTRNDVGP YECGIQNELS VDHSDPVILN VLYGDDPTI 420
SPSYTYRPG VNLSSLCHAA SNPPAQYSWL IDGNIQHQHT ELFISNITEK NSGLYTCQAN 480
NSAGHSRRT VKTITVSAEL PKPSSISNNS KPVEDKDAVA FTCEPEAQNT TYLWVWNGQS 540
LPVSPRLQLS NGRNTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDLV YGPDPTIISP 600
PDSSYLSGAN LNLCHSASN PSPQYSWRIN GIPOQHTQVL FIAKITPNNN GTYACFVSNL 660
ATGRNNSIVK SITVSASGTS PGLSAGATVG IMIGVLVGVA LI

```

Seq ID NO: 534 DNA sequence  
Nucleic Acid Accession #: NM\_006952.1  
Coding sequence: 11..793

65  
70  
75  
80

```

1 11 21 31 41 51
| | | | |
AATCCCGACA ATGGCGAAGG ACAACTCAAC TGTTCGTTGC TTCCAGGGCC TGCTGATTTT 60
TGGAAATGTG ATATTGTGTT GTTGGCGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120
ATCTGACCAA CACAGCCTCT ACCCACTGCT TGAAGCCACC GACAACGATG ACATCTATGG 180
GGCTGCTTGG ATCGCATATG TTGTGGGCAT CTGCTCTTTC TGCTGTCTGT TTCTAGGCAT 240
TGTAGGCATC ATGAAGTCCA GCAGGAAAAT TCTTCTGGCG TATTTCAITC TGATGTTTAT 300
AGTATATGCC TTTGAAGTGG CATCTTGAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360
ACCAACCTCT TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAGCC CTCCAAACAA 420
TGATGACCAG TGGAAAAACA ATGGAGTCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480
CAATGTCTGT GCGGTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC 540
TGAGAATAAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600
AGAACCTCTC AACCTGGAGG CTGTGAAACT AGGCGTGCCT GGTTTTATC ACAATCAGGG 660
CTGTATGAA CTGATCTCTG GTCCAATGAA CCGACACGCC TGGGGGGTTG CCGTGTGTTG 720
ATTGTCCATT CTGTCTGGA CTTTTGGGT TCTCTGGGT ACCATGTCTT ACTGGAGCAG 780
AATTGAATAT TAAGAA

```

Seq ID NO: 535 Protein sequence  
Protein Accession #: NP\_008883.1

85

```

1 11 21 31 41 51
| | | | |

```

MAKDNSTVRC FQGLLIFGNV IIGCCGIALT AECIFFVSDQ HSLYPLLEAT DNDDIYGAAN 60  
 IGFVGLICLP CLSVLGIIVGI MKSSSRKILLA YFILMFIVYA FEVASCITAA TQRDFFTPNL 120  
 FLKQMLEERYQ NNSPPNNDDQ WKNNGVTKTW DRLMLQDNCC GVNGPSDWQK YTSAFRTENN 180  
 DADYWPWRQC CVMNNLKBPQ NLEACKLGVP GFYHNQGCYE LISGPMNRHA WGVANWFGEAI 240  
 LCWTFWVLLG TPFYWSRIEY

Seq ID NO: 536 DNA sequence  
 Nucleic Acid Accession #: NM\_002638.1  
 Coding sequence: 120..473

1 11 21 31 41 51  
 CAATACAGCT AAGGAATTAT CCCTTGTAAT TACCACAGAC CCGCCCTGGA GCCAGGCCAA 60  
 GCTGGACTGC ATAAAGATTG GTATGGCCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120  
 TGAGGGCCAG CAGCTTCTTG ATCGTGGTGG TGTTCTCATC CGCTGGGACG CTGGTTCTAG 180  
 AGGCAGCTGT CACGGGAGTT CCTGTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCAT 240  
 TCAATGGACA AGATCCCGTT AAAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300  
 CGCAAGAGCC AGTCAAAGGT CCACTCTCCA CTAAGCCTGG CTCCTGCCCC ATTATCTTGA 360  
 TCCGTTGGCG CATGTGAAT CCCTTAACC GCTGCTTGA AGATACTGAC TGCCAGGAA 420  
 TCAAGAAAGT CTGTGAAGGC TCTTGGGGA TGGCTGTGTT CGTCTCCAG TGAAGGGAGC 480  
 CGGTCTTGGC TGCACTGTG CCGTCCCGAG AGCTACAGG CCCATCTGGT CCTAAGTCCC 540  
 TGCTGCCCTT CCGCTTCCA CACTGTCCAT TCTTCTCCC ATTCAGGATG CCCAGGCTG 600  
 GAGCTGCTC TCTCATCCAC TTCCAATAA A

Seq ID NO: 537 Protein sequence  
 Protein Accession #: NP\_002629.1

1 11 21 31 41 51  
 MRASSFLIVV VFLIAGTLVL EAAVTGVPVK GQDTVGRVP FNGQDPVKQ VSVKGQDKVK 60  
 AQEPVKGPVS TKPGSCPIIL IRCAMLNPN RCLKDTDCPG IKKCEGSCG MACFPVQ

Seq ID NO: 538 DNA sequence  
 Nucleic Acid Accession #: NM\_001793.2  
 Coding sequence: 71..2560

1 11 21 31 41 51  
 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CCGCCGTCGC GGCAGCTGCT TCACCCCTCT 60  
 CTCTGAGCC ATGGGGCTCC CTCGTGGACC TCTCGGCTCT CTCTCTCTTC TCCAGGTTTG 120  
 CTGGCTGCGA TGCGCGGCTC CCGAGCCGTG CCGGGCGGTG TTCAGGGAGG CTGAAGTGAC 180  
 CTTGGAGCGG GGAGGCGCGG AGCAGGAGCC CGGCCAGGCG CTGGGGAAGG TATTATGGG 240  
 CTGCCCTGGG CAAGAGCCAG CTCGTGTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300  
 TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAGG AATCCATTGA AGATCTTCCC 360  
 ATCCAAACCT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCAA TATCTGTCCC 420  
 TGAATAATGC AAGGGTCCCT TCCCCAGAG ACTGAATCAG CTCAGTCTA ATAAAGATAG 480  
 AGACACCAAG ATTTTCTACA GCATCAGCGG CCGCGGGGCA GACAGCCCCC CTGAGGGTGT 540  
 CTTGCTGTGA GAGAGGAGA CAGGCTGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600  
 GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTGAGG AATGGTGCCT CAGTGGAGGA 660  
 CCCCATGAAC ATCTCCATCA TGTGACCGA CCAAGATGAC CACAAGCCCA AGTTTACCCA 720  
 GGACACCTTC CGAGGAGATG TCTTAGAGGG AGTCTTACCA GGTACTTCTG TGATGCAGGT 780  
 GACAGCCAGC GATGAGGATG ATGCCATCTA CACTACAAT GGGGTGGTGT CTTACTCCAT 840  
 CCATAGCCAA GAACCAAGAG ACCCAACGGA CCTCATGTTT ACCATTCAAC GGAGCACAGG 900  
 CACCATCAGC GTCATCTCCA GTGGCCTGGA CCGGGAAGAA GTCCCTGAGT ACACACTGAC 960  
 CATCCAGGCC ACAGACATGG ATGGGGAAGG CTCACCAACC ACGGCAGTGG CAGTAGTGGG 1020  
 GATCCTTGAT GCCAATGACA ATGCTCCCAT GTTTGACCCC CAGAAGTACG AGGCCCATGT 1080  
 GCCTGAGAAT GCGTGGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140  
 CAATCAACA CGGTGGGCTG CCACCTACCT TATCATGGGC GGTGAAGACG GGGACCATTT 1200  
 TACCATCACC ACCCACCTTG AGAGCAACCA GGGCATCCTG ACACACGAGG AGGGTTTGGG 1260  
 TTTTGAAGCC AAAAACCCAG ACACCTGTA CGTTGAAGTG ACCAAGAGG CCCCTTTTGT 1320  
 GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGTCCAG GTGGAGGATG TGAATGAGGC 1380  
 ACCTGTGTTT GTCCACCTC CCAAGTCTG TGAGGTCCAG GAGGCGATCC CCACTGGGGA 1440  
 GCCTGTGTGT GTCTACACTG CAGAAGACCC TGACAAGGAG AATCAAAAGA TCAGCTACCG 1500  
 CATCTGAGA GACCCAGCAG GGTGGCTAGC CATGGACCCA GACAGTGGGC AGGTACAGC 1560  
 TGTGGGCACC CTCGACCGTG AGGATGAGCA GTTTGTGAGG AACACATCT ATGAAGTCAT 1620  
 GGTCTTGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGTAAAC 1680  
 ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGCCC CGTCAGATCA CCATCTGCRA 1740  
 CCAAGCCCT GTGGCCAGG TGCTGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800  
 CCCTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860  
 GGAAGGTGAC ACAGTGTCT GTTCCCTGAA GAAGTTCTCT AAGCAGGATA CATATGACGT 1920  
 GCACCTTTCT CTGTGTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980  
 GTGGCAGTGC CATGGCCATG TCGAAACCTG CCCTGGACCC TGGAGGGAGG GTTTCATCCT 2040  
 CCCTGTGCTG GGGGCTGTCC TGGCTCTGCT GTTCTCTCTG CTGGTGTCTG TTTTGTGTGT 2100  
 GAGAAAGAG CGGAGATCA AGGAGCCCTC CTTACTCCCA GAAGATGACA CCGGTGACAA 2160  
 CGTCTTCTAC TATGGCCAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220  
 GCTCCACCGA GGTCTGAGG CAGGCGGGA GGTGGTCTC CGCATGAGC TGGCACCAC 2280  
 CATCATCCCG ACACCCATGT ACCGTCTCTG GCCAGCCAC CCAAGTGAAG TCGGCAACTT 2340  
 TATAATTGAG AACCTGAGG CGGCTAACAC AGACCCCA CA GCGCCGCTC ACAGACCCCT 2400  
 CTGTGTGCTG GACTATGAGG CGACGGCTC CGACGCGCG TCCCTGAGCT CCCTCACCTC 2460  
 CTCGCTCTCC GACCTGCTAC AGATTACGA TTATCTGAAC GAGTGGGGA CCGGCTTCAA 2520  
 GAAGCTGGCA GACATGTACG GTGGCGGGA GGACGACTAG CGGCTCTGCC TGCAGGGCTG 2580  
 GGGACCAAC GTCAAGCCAC AGAGCATCTC CAAGGGCTCT CAGTTCCCCC TTCAGCTGAG 2640  
 GACTTCGGAG CTGTGAGGA AGTGGCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700  
 ACGTTAGAGT GGTGTCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTGTACTTC 2760  
 AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCCTCAG AGGCCAAGT TCCAGAGGCC 2820  
 TCTTACCTGC CGTAAATGC TCAACCTGT GTCTGGGCC TGGGCTGCT GTGACTGACC 2880  
 TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGGTGCA ACTTAATTTT 2940

TTTTTTTAAAT GCTATCTTCA AAACGTTAGA GAAAGTTCCT CAAAAGTGCA GCCCAGAGCT 3000  
 GCTGGGCCCA CTGGCCGCTCC TGCAATTCCTG GTTTCAGAC CCCAATGCCT CCCATTCGGA 3060  
 TGGATCTCTG CGTTTTTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTT TATTTCCCT 3120  
 GTTGGCTTGC TATAGATGAA GGGTGAGGAC AATCGTGAT ATGTACTAGA ACTTTTTTAT 3180  
 TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 539 Protein sequence  
 Protein Accession #: NP\_001784.2

1 11 21 31 41 51  
 MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG 60  
 QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120  
 KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREBIK 180  
 YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240  
 DEDDAIYTYN GVWAYSIHSQ EPKDPEDLMP TIHRSTGTIS VISSGLDREK VPEYTLTIQA 300  
 TDMGDGSGTT TAVAVVEILD ANDNAPMFDQ QKYEAHVPEP AVGHVQRLT VTDLDAPNSP 360  
 ANRATYLIMG DGDGDHFTIT THPESNQIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420  
 PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480  
 DPAGWLAMP DSGQVTAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLTLID 540  
 VNDHGVPPEP RQITICNQSP VRQVLNITDK DLSPHTSPPQ AQLTDDSDIY WTAEVNEEGD 600  
 TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFILFVL 660  
 GAVLALLFL LVLLLLVRK RKIKEPLLLP EDDTRDNVFP YGEGGGED QDYDITQLHR 720  
 GLEARPEVVL RNDVAPTIIIP TPHYRPRPAN PDEIGNPIIE NLKAANTDPT APPYDTLLVF 780  
 DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

Seq ID NO: 540 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..672

1 11 21 31 41 51  
 ATGAGGCTCC AAAGACCCCG ACAGGCCCGG CGGGGTGGGA GCGCGCGGCC CGGGGCGGG 60  
 CGGGGCTCCC CCTACCGGCC AGACCGCGGG AGAGGCGGCG GAGGCTGCG AAGGTTCCAG 120  
 AAGGCGGGG AGGGGCGGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG 180  
 CTGCTCGCCT TCGCTGCTGT TCGGCGCTA CCGCGGCTGT GGACGAGCG CAACCTGACT 240  
 GCGAGACAC GAGATCCAGA GGAATCCAG CGAACGAGCG AGGCTGACAA TAGAGTGTGG 300  
 TGTCACTGTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAATGG 360  
 ACAGAGCCAT ACTGCGTAT AGCGCGCGTG AAAATATTTT CAGCTTTTTT CATGTTGCG 420  
 AAGCAGTGCT CCGCTGCTGT TGCAGCGATG GAGAGACCCA AGCCAGAGGA GAAGCGGTTT 480  
 CTCTCGAAG AGCCATGCGC CTCTTTTAC CTCAAGTGT GTAAATTCG CTACTGCAAT 540  
 TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAGAAT ATGCTGGGAG CATGGGTGAG 600  
 AGCTGTGTG AGCTGTGCT GGCATCTCT CTGCTGCTGG CTCTCATTC AGCGGCGCTC 660  
 AGCCTGTCTT GA

Seq ID NO: 541 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MRLQRPRQAP AGGRRAPRGG RGSFYRPDPG RGARRLRFPQ KGGEGAPRAD PFWAPLGTMA 60  
 LLALLLVVAL PRVWTDANLT ARQRPEDSQ RTDEGDNVWV CHVCEENTP EQCNPRCKW 120  
 TEPYCVIAAV KIPFRFFMVA KQCSAGCAAM ERKPEEKRF LLEPMPPFY LKCKKIRYCN 180  
 LBGPPIINSSV FKEYAGSMGE SCGLWLAAIL LLLASIAAGL SLS

Seq ID NO: 542 DNA sequence  
 Nucleic Acid Accession #: XM\_035292.2  
 Coding sequence: 53..1576

1 11 21 31 41 51  
 GCTCGCTGGG CGCGGCTCC CGGGTGTCCC AGGCCCGGCC GGTGCGCAGA GCATGGCGGG 60  
 TGCGGGCCCG AAGCGGCGCG CGCTAGCGGC GCGGCGCGCC GAGGAGAAGG AAGAGGCGCG 120  
 GGAGAAGATG CTGGCCGCCA AGAGCGCGGA CGGCTCGGCG CCGGCGAGCG AGGGCGAGGG 180  
 CGTGACCCCTG CAGCGGAACA TCACGCTGCT CAACGCGGTG GCCATCATCG TGGGGAACAT 240  
 TATCGGCTCG GGCATCTTCG TGAAGCCAC GGGCGTGTCT AAGGAGGCG GCTCGCGGG 300  
 CTGCGGCTCG GTGGTGTGGG CCGGTGTGGG CGTCTTCTCC ATCGTGGCG CGCTCTGCTA 360  
 CGCGGAGCTC GGCACACCA TCTCAAAATC GGGCGGCGAC TACGCTACA TGCTGGAGGT 420  
 CTACGCTCG CTGCGCGCTT TCTCAAGCT CTGGATCGAG CTGCTCATCA TCGGCGCTTC 480  
 ATCGCAGTAC ATCGTGGCCC TGGTCTTGGC CACCTACCTG CTCAAGCGCG TCTTCCCGAC 540  
 CTGCGCGGTG CCGGAGGAG CAGCGAAGCT CGTGCGCTGC CTCTGGTGC TGCTGCTCAC 600  
 GGCGGTGAAC TGCTACAGCG TGAAGCGCG CACCGCGGTC CAGGATGCCT TTGCGCGCGC 660  
 CAAGCTCTCG GCGTGGCCC TGATCATCTT GCTGGGCTTC GTCCAGATCG GAAAGGGTGA 720  
 TGTGTCCAAAT CTAGATCCCA ACTTCTCATT TGAAGGCACC AAACCTGGATG TGGGGAACAT 780  
 TGTGCTGGCA TTATACAGCG GCGCTCTTGC CTATGGAGGA TGGAAATTA TGAATTTCTG 840  
 CACAGAGGAA ATGATCAACC CTACAGAAA CCGTCCCGTG GCCATCATCA TCTCCCTGCC 900  
 CATCGTGACG CTGGGTGACG TGCTGACCAA CCGTGCCTAC TTCACACCC TGTCACCGA 960  
 GCAGATGCTG TCGTCCGAGG CGTGGCGGT GGAATTCGCG AACTATCACC TGGGCTCAT 1020  
 GTCCTGGATC ATCCCGCTCT TCGTGGCGCT GTCCTGCTTC GGCTCCGTC ATGGGTCCCT 1080  
 GTTCATATCC TCCAGGCTCT TCTTGTGGG GTCCCGGAA GGCACCTGC CCTCATCT 1140  
 CTCCATGATC CACCCACAGC TCTCACCCC CGTGGCGTCC CTGCTGTTC CGTGTGTGAT 1200  
 GACGCTGCTC TAGGCTTCT CCAAGGACAT CTCTCGGTG ATCAACTTCT TCAGCTTCT 1260  
 CAACTGGCTC TGGGTGGCCC TGGCCATCAT CGGCATGATC TGGCTGCGCC ACAGAAAGCC 1320  
 TGAGCTTGAG CGGCCATCA AGGTGAACCT GGCCTGCGCT GTGTTCTTCA TCGTGGCTG 1380  
 CTTCTTCTG CTGCGGCTCT CTTCTGGAA GACACCGTG GAGTGTGGCA TCGGCTTCA 1440  
 CATCATCTCT AGCGGCTGC CGTCTACTT CTCTGGGTC TGGTGGAAA ACAAGCCCAA 1500  
 GTGGCTCTC CAGGCACTCT TCTCACGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGTT 1560



CCCCCAGGAG ACATAGCCAG GAGGCCGAGT GGCTGCCGGA GGAGCATGC

Seq ID NO: 543 Protein sequence  
Protein Accession #: XP\_035292.2

5  
10  
15  
1  
11 21 31 41 51  
MAGAGPKRRA LAAPAAEKEE EAREKMLAAK SADGSAPAGE GEGVTIQRNI TLLNGVAIIV 60  
GTIIGSGIFV TPTGVLKEAG SPGLALVVWA ACGVFSIVGA LCYAEIGTTI SKSGGDYAYM 120  
LEVYGSIPAF LKLWIELLII RPSSQYIVAL VFATYLLKPL FPTCPVPEEA AKLVACLCLV 180  
LLTAVNCYSV KAATRVQDAF AAKLLALAL IILIGFVQIG KGDVSNLDPN FSFEGTKLDV 240  
GNIVLALYSG LFAYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVYV LTNLAYFTTL 300  
STEQMLSEEA VAVDFGNVHL GVMSWII PVF VGLSCFGSVN GSLFTSSRLP FVGSREGHLP 360  
SILSMIHPQL LTPVPSLVFT CVMTLLYAFS KDIFSVINFP SPFNWLCVAL AIIGMIWLRH 420  
RKPELERPIK VNLALPVFFI LACLFLIAVS FWKTPVECGI GPTIILSGLP VYFFGVWVWN 480  
KPKWLLQGIF STTVLCQKLM QVVPQET

Seq ID NO: 544 DNA sequence  
Nucleic Acid Accession #: NM\_005268.1  
Coding sequence: 168..989

20  
25  
30  
35  
40  
45  
1  
11 21 31 41 51  
TAAAGAGCAA AAGAATTCGC GGGCGGCTCG ACACGGGCTT CCGCGAAAC CTTCCTCCGCT 60  
TCTGGATATG AAATTCACAG TGCTTGCTGA GTCTTATGCG CGGCTGTCTGG GAGCCAGGAG 120  
AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACGCGTGG GTCCACCATG AACTGGAGTA 180  
TCTTTGAGGG ACTCCTGAGT GGGGTCAACA AGTACTCCAC AGCCTTTGGG CGCATCTGGC 240  
TGCTCTGGT CTTCATCTTC CGCTGCTGG TGTACCTGCT GAGCGCCGAG CGTGTCTGGA 300  
GTGATGACCA CAAGGACTTC GACTGCAATA CTCGCCAGCC CGGCTGCTCC AACGTCCTGCT 360  
TTGATGAGTT CTTCCTCTGT TCCCATGTGC GCCTCTGGGC CCTGCACTT ATCCTGTGTA 420  
CATGCCCTCT ACTGCTGTGT GTCATGCACG TGGCCTACCG GGAGGTTTCA GAGAAGAGGC 480  
ACCGAGAAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG 540  
GTGGGCTCTG GTGGACATAT GTCTGCAGCC TAGTGTTCAC GGCAGAGCTG GACATCGCCT 600  
TTCTCTATGT GTTCCACTCA TTCTACCCCA AATATATCTT CCTCTCTGTG GTCAAGTGCC 660  
ACGAGATACC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAAGCCCTCA GAGAAGAACA 720  
TTTTCACCTT CTTCATGGTG GCCACAGCTG CCATCTGCAT CCTGCTCAAC CTCGTGGAGC 780  
TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCTTGGC AGCAAGGAAA GCTCAGGCCA 840  
TGTGCACAGG TCATCACCCC CACGGTACCA CCTCTCTCTG CAAACAAGAC GACCTCCTTT 900  
CGGGTGACCT CATCTTTCTG GGCTCAGACA GTCATCTCTC TCTCTTACCA GACCGCCCCC 960  
GAGACCATGT GAAGAAACAC ATCTTGTGAG GGGCTGCTCT GACTGTGCTG GCAGGTGGGG 1020  
CCTGGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC 1080  
CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGACGCTC TGGGAGCCAG TTCCTAGTCC 1140  
TCAACTCCAG CCACCTGCCC CAGCTCGACG GCACCTGGCC AGTTCCTCCCT CTGCTCTGCA 1200  
GCTCGGTTTC CTTTCTAGA ATGGAATAG TGAGGGCCAA TGC

Seq ID NO: 545 Protein sequence  
Protein Accession #: NP\_005259.1

50  
55  
1  
11 21 31 41 51  
MNWSIFEGLL SGVKNYSTAF GRINLSLVFI FRVLVYLVT A ERVMSDDHDK PDCNTRQPGC 60  
SNVCFDEFFP VSHVRLNALQ LILVTCPSLL VVMHVAYREV QEKRRHREAH ENSGRLYLNP 120  
GKKRGLWMT VVCSLVFKAS VDIAPLYVFH SFYPKYILPP VVKCHADPCP NIVDCFISKP 180  
SEKNIFTLFM VATAAICILL NLVELIYLV S KRCHECLAAR KAQAMCTGHH PHGTTSSCKQ 240  
DDLSSGLDIF LGSDSHPPLL PDRPRDHVK TIL

Seq ID NO: 546 DNA sequence  
Nucleic Acid Accession #: NM\_002391.1  
Coding sequence: 26..457

60  
65  
70  
75  
1  
11 21 31 41 51  
CGGGGGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCACCTCTCT 60  
CGCCCTGCTG GCGCTCACTT CCGCGGTGCG CAAAAGAGAA GATAAGGTGA AGAAGGGCGG 120  
CCGGGGGAGC GAGTGCCTG AGTGGGCTTG GGGGCCCTGC ACCCCAGCA GCAAGGATTG 180  
CGGCGTGGGT TTCCGCGAGG GCACCTGCGG GGCCAGACC CAGCGCATCC GGTGCAAGGT 240  
GCCCTGCAAC TGAAGAAGG AGTTTGGAGC CGACTGCAAG TACAAGTTTG AGAACTGGGG 300  
TGGGTGTGAT GGGGGCACAG GCACCAAGT CCGCCAAGGC ACCCTGAAGA AGGCGCGCTA 360  
CAATGCTCAG TGCCAGGAGA CCATCGCGT CACCAAGCCC TGCACCCCA AGACCAAGC 420  
AAAGGCCAAA GCCAAGAAAG GGAAGGGAAG GGAAGGGAAG CCAAGCCTGG ATGCCAAGGA 480  
GCCCTGGTGT TCACATGGGG CCTGGCCACG CCTCTCTCT CCCAGGCCCG AGATGTGACC 540  
CACCAGTGCC TTCTGTCTGC TGTGTAGCTT TAATCAATCA TGCCCTGCCT TGTCTCTCTC 600  
ACTCCCGAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT 660  
TGAGCTTCCC CCAAGCAAT GTGAGTCCCA GAGCCGCTT TTGTTCTTCC CCACAAATCC 720  
ATTACTAAGA AACACATCAA ATAACTGAC TTTTCCCCC CAATAAAGC TCTTCTTTT 780  
TAATAT

Seq ID NO: 547 Protein sequence  
Protein Accession #: NP\_002382.1

80  
85  
1  
11 21 31 41 51  
MQHRGFLLLT LLALLAL TSA VAKKDKVKK GPGSECAEW AWGPCTPSSK DCGVGFREGT 60  
CGAQTORIRC RVPCNWKKEF GADCKYKFN WGACDGGTGT KVRQGTLLKA RYNAQCQETI 120  
RVTKPCTFKT KAKAKAKGK GKD

Seq ID NO: 548 DNA sequence



Nucleic Acid Accession #: NM\_006783.1

Coding sequence: 1..786

5 1 11 21 31 41 51  
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAAACACTC CACCAGCATC 60  
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCGAGTCA TGATCCTAGT GGTGGCTGCC 120  
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCTCTGCA ACACACTGCA ACCGGGATGC 180  
10 AAAAAATGTGT GCTATGACCA CTTTTCCTCG GTGTCCACA TCGGCTGTG GGCCTCCAG 240  
CTGATCTTCTG TCTCCACCCC AGCGCTGCTG GTGGCCATGC ATGTGGCCTA CTACAGGCAC 300  
GAAACCACTC GCAAGTTCAG GCGAGGAGAG AAGAGGAATG ATTTCAAAGA CATAGAGGAC 360  
ATTAAAAAGC ACAAGTTCG GATAGAGGGG TCGCTGTGTT GGACGTACAC CAGCAGCATC 420  
TTTTTCCGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACATGGG 480  
15 TACCACCTGC CTTGGGTGTT GAAATGTGGG ATTGACCCCT GCCCAACCT TGTGACTGC 540  
TTTATTTCTA GGCCAAACAGA GAAGACCGTG TTTACCATTT TATGATTTC TGCGTCTGTG 600  
ATTGTGATGC TGCTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTATGG 660  
AGATCAAAGA GAGCACAGAC GCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720  
CAGAAATGAAA TGAATGAGCT GATTTTCAAT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780  
AGCTAA

Seq ID NO: 549 Protein sequence

Protein Accession #: NP\_006774.1

25 1 11 21 31 41 51  
MDWGLTHTFI GGVNKHSTSI GKVNITVIFI FRVMILVVA QEVWGEDEQED FVCNTLQPGC 60  
KNVCYDHPFP VSHIRLWALQ LIPVSTPALL VAMEVAYYRH ETTRKFRRGE KRNDPKDIED 120  
IKHKVRVIEG SLWMTYTSI PFRIIFEAAP MYVPVFLYNG YHLPWVLKCG IDPCPNLVDC 180  
30 FISRPTEKTV PTIFMISASV ICMLLNVAEL CYLLLKVCFR RSKRAQTQKN HPNHALKESK 240  
QNMENELISD SGQNAITGFP S

Seq ID NO: 550 DNA sequence

Nucleic Acid Accession #: NM\_002571.1

Coding sequence: 99..587

35 1 11 21 31 41 51  
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCGC AGCCATGCTG TGCCCTCTGC 60  
TCACCTCTGG CGTGGCCCTG GTCTGTGGTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120  
40 AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CCTGGCACTC CATGGCCATG GCGACCAACA 180  
ACATCTCCCT CATGGCGACA CTGAAGGCCCT CTCTGAGGGT CCACATCACC TCACTGTTGC 240  
CCACCCCCGA GGACAACTCT GAGATCGTTC TGACAGATG GGAGAACAAC AGCTGTGTTG 300  
AGAAGAAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACGG 360  
45 TGGCGAAGCA GGCACGCTG CTGATACCTG ACTACGACAA TTCTCTGTTT CTCTGCCTAC 420  
AGGACACCCAC CACCCCCATC CAGAGCATGA TGTGCCAGTA CCTGGCCAGA GTCCTGGTGG 480  
AGGACGATGA GATCATGAGG GGATTCATCA GGGCTTTGAG GCCCTTGCCC AGGCACCTAT 540  
GGTACTTGCT GGACTTGAAG CAGATGGAAG AGCGGTGCGG TTTCTAGCTC ACCTCCGCTT 600  
CCAGGAAGAC TCCGAGGAG CCCTTCCACA CCTCCAGAGC AGTGGGACTT CCTCTGAGCC 660  
50 TTTCAGAGAA TAACACAGC TCAGAAGACG ATGAGTGGT CATCTGTGTC GCCATCCCTT 720  
TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCTGGGGGC AGAGTCTCTG 780  
GCAGAGGTTA TTAATAAACCC CTTGAGCAT G

Seq ID NO: 551 Protein sequence

Protein Accession #: NP\_002562.1

55 1 11 21 31 41 51  
MDIPQTKQDL ELPLKAGTWH SMAMATNNIS LMATLKAPLR VHITSLPTP EDNLEIVLHR 60  
60 WENNSCIVEKK VLGEKTNPK KFKINYTVAN EATLLDSTDYD NFLFLCLQDT TPIQSMCMQ 120  
YLARVLVEDD EIMQGFIRAF RPLPRHLNWL LDLQMEEPK RF

Seq ID NO: 552 DNA sequence

Nucleic Acid Accession #: NM\_006500.1

Coding sequence: 27..1967

65 1 11 21 31 41 51  
ACTTGGTCTC GCGCTCCCGG CCAAGCATGG GGCTTCCCAG GCTGGTCTGC GCCTTCTTGC 60  
70 TCGCCGCTCT CTGCTGCTGT CCTCGGCTGG CGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120  
CGCTGAGCT GGTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCCTCTCCC 180  
AGTCCCAAGG CAACCTCAGC CATGTGACT GTTTTCTGT CCACAAGGAG AAGCGGACGC 240  
TCATCTTCCG TGTGGCCGAG GGCAGGGGCC AGAGCGAACC TGGGAGTAC GAGCAGCGGC 300  
TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCTGACTCA AGTCACCCCC CAAGACGAGC 360  
75 GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCCGCG 420  
TCTACAAAGC TCCGAGGAG CCAACATCC AGGTCAACCC CCTGGGCATC CCTGTGAACA 480  
GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGGTACCCC ATTCTCTAAG 540  
TCATCTGGTA CAAGAAATGGC CGGCCTCTGA AGGAGGAGAA GAACCGGCTC CACATTCACT 600  
CGTCCAGAGC TGTGGAGTGG AGTGGTTTGT ACACCTTGCA GAGTATCTG AAGGCACAGC 660  
80 TGTGTAAGA AGACAAAGAT GCCCAGTTT ACTGTGAGCT CAACCTACCG CTGCCAGTG 720  
GGAACCAATC GAGGAGGTC AGGGAGTCA CGTCCCTGT TTTCTACCG ACAGAAAAGG 780  
TGTGGCTGGA AGTGAGGCC GTGGGAATGC TGAAGGAAGG GAACCGGCTG GAAATCAGGT 840  
GTTTGGCTGA TGGCAACCTT CCACCACTC TCAGCATCAG CAAGCAGAAC CCCAGACCA 900  
GGGAGGCAGA GGAAGAGACA ACCAAGACA ACGGGTCTCT GGTGCTGGAG CCTGCCCGGA 960  
AGGAACACAG TGGCGGCTAT GAATGTCAAG CCTGGAACCT GGACACCATG ATATCGCTGC 1020  
85 TAGATGAACC ACAGGAACCTA CTGGTGAAT ATGTGTCTGA CGTCCGAGTG AGTCCCGCAG 1080  
CCCTGAGAG ACAGGAAGGC AGCAGCCTCA CCTGACCTG TAGGAGCAGG AGTAGCCAGG 1140  
ACCTCGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAGGT GCTGGAAGG GGCCTGTGTC 1200

5 TTAGCTTGA TGAACCTGAAA CGGAGAGCAG GAGGCGGCTA TCGCTGCGTG GCGTCTGTGC 1260  
 CCAGCATACC CGGCTGAAAC CGCACACAGC TGGTCAAGCT GGCCATTTT GGCCCCCCTT 1320  
 GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAGAGAA TATGGTGTG AATCTGTCTT 1380  
 GTGAAGCGTC AGGGCACCCC CGGCCACCCA TCTCTGGAA CGTCAACGGC ACGGCAAGTG 1440  
 AACAGAGCCA AGATCCACAG CGAGTCTCTGA GCACCTCGAA TGCTCTCGTG ACCCCGGAGC 1500  
 TGTGTGAGAC AGGTGTGAA TGAACGGCCT CCAACGACCT GGGCAAAAC ACCAGCATCC 1560  
 TCTTCTGGA GCTGGTCAAT TTAACACCCC TCACACCAGA CTCACACACA ACCACTGGCC 1620  
 TCAGCACTTC CACTGCCAGT CCTCATACCA GAGCCAACAG CACCTCCACA GAGAGAAAGC 1680  
 10 TGCCGGAGCC GGAGAGCCGG GCGTGGTCA TCGTGGCTGT GATTGTGTGC ATCTGGTGC 1740  
 TGGCGGTGCT GGGCGCTGTC CTCTATTTC TCTATAAGAA GGGCAAGCTG CCGTGCAGGC 1800  
 GCTCAGGGAA GCAGGAGATC ACCTGCCCCC CGTCTCGTAA GACCGAACTT GTAGTTGAAG 1860  
 TTAAGTCAGA TAAGTCTCCA GAAGAGATGG GCCTCTGCA GGGCAGCAGC GGTGACAAAG 1920  
 GGGCTCCGGG AGACAGAGGA GAGAAATACA TCGATCTGAG GCATTAGCCC CGAATCACTT 1980  
 15 CAGCTCCCTT CCGTGCCTGG ACCATTCCCA GCTCCCTGCT CACTCTTCTC TCAGCCAAAG 2040  
 CCTCCAAAGG GACTAGAGAG AAGCCTCTCT CTCCCTCTAC CTGCACACCC CCTTTCAGAG 2100  
 GGCCACTGGG TTAGGACCTG AGGACCTCAC TTGGCCCTGC AAGCCGCTTT TCAGGGACCA 2160  
 GTCCACCACC ATCTCTCCCA CGTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCCAGTCTC 2220  
 CCGAGCGGTG AGGAGAGTGT CTTCAGAAC GTGTTTTTTC TTACACACA TTATGGCTGT 2280  
 20 AAATACCTGG CTCTGCCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTGG TTTCTGCC 2340  
 CAAGAGCTGG CTTCACCAT CCAGGTGCAC CACTGAAGTG AGGACACACC GGAGCCAGGC 2400  
 GCGTGTCTAT GTTGAAGTGC GCTGTTTACA CCGCTCCGG AGAGCACCCC AGCGGCATCC 2460  
 AGAAGCAGCT GCAGTGTGTC TGCCACCACC CTCTCTCG CTCTTCAA GTCTCTGTG 2520  
 ACATTTTTTC TTTGGTCAGA AGCCAGAAC TGGTGTCTT CTTAAAGA TACGTGCCGG 2580  
 25 GGCCAGGTGT GGTGGCTCAC GCCTGTATC CCAGCACTT GGGAGGCCGA GGGGGGCCGA 2640  
 TCACAAAGTC AGGACAGAGC CATCCTGGCT AACACGGTGA AACCTGTCT CTACTAAAA 2700  
 TACAAAAAAA AATTAGCTAG GCGTAGTGGT TGGCACCTAT AGTCCAGCT ACTCGGAAGG 2760  
 CTGAAGCAGG AGAATGGTAT GAATCCAGGA GGTGGAGCTT GCAGTGAGCC GAGACCGTGC 2820  
 CACTGCACCT CAGCCTGGGC AACACAGCGA GACTCGTCT CGAGGAAAAA AAAAGAAAAA 2880  
 30 ACGGTACCT GCGGTGAGGA AGCTGGGCGC TGTTCCTGAG TTCAGGTGAA TTAGCTCAA 2940  
 TCCCGGTGTT CACTTGTCTC CATAGCCCTC TTGATGATC ACGTAAACT GAAAGGCAGC 3000  
 GGGGAGCAGA CAAGATGAG GTCTACACTG TCCTTCATGG GGATTAAGC TATGTTTATA 3060  
 TTAGCACCAG ACTTCTACAA ACCAAGCTCA GGGCCCCAAC CCTAGAAGGG CCAAAATGAG 3120  
 AGAATGGTAT TTAGGGATGG AAAACGGGGC CTGGCTAGAG CTTCGGGTGT GTGTGTCTGT 3180  
 35 CTGTGTGTAT GCATACATAT GTGTGTATAT ATGTTTTTGT CAGGTGTGTA AATTGCAAA 3240  
 TTGTTTCTCT TATATATGTA TGTATATATA TATATGAAA TATATATATA TATGAAAAAT 3300  
 AAAGCTTAAT TGTCCAGAA AATCATACAT TGCTTTTTTA TTCTACATGG GTACACAGG 3360  
 AACCTGGGG CCTGTGAAC TACAACCAA AGGCACCAA AACCGTTTCC AGTTGGCAGC 3420  
 40 AGAGATCAGG GGTACCTCT GCTTCTGAGC AAATGGCTCA AGCTCTACCA GAGCAGACAG 3480  
 CTACCTACT TTTAGCAGC AAAACGTCCC GTATGACGCA GCACGAAGGG CTTGGCAGGC 3540  
 TGTAGCAGG AGCTATGTCC CTTCCTATCG TTTCCGTCCA CTT

Seq ID NO: 553 Protein sequence  
Protein Accession #: NP\_006491.1

45 1 11 21 31 41 51  
 GLPRLVCAFL LAACCCPRV AGVPGAEQF APELVEVEVG STALLKGLS QSQGNLSHVD 60  
 WFSVHKKRT LIFRVROGG QSEPGYEQR LSLQDRGATL ALTQVTPQDE RIFLCQGRP 120  
 50 RSQBYRIQLR VYKAPBEPNI QVNPLGIPVN SKEPEEVATC VGRNGYPIPO VIWYKNGRPL 180  
 KEENRNVHIQ SSQTVSSSL YTLQSLKQ LKEDKDAQF YCELNYRLPS GNMKESREV 240  
 TVPVFYPTFK WLEVEVPVGM LKEGDRVEIR CLADGNPPPH PSISKQNPST REAEETIND 300  
 NGVLVLEPAR KEHSGRYEQ ANWLDYMSL LSEPELNV YVSDVRVSPA APERQEGSSL 360  
 TLTCEAESSQ DLEFQWLRBE TDQLERGPV LQLHDLKREA GGGYRCVASV PSIPGLNRTQ 420  
 55 LVKLAIFGPP WNAPEKRVV VKENMVLNLS CEASGHPRPT ISWNVNGTAS EQDQDPQVRL 480  
 STLNLVLTPE LLETGVBCTA SMDLGNKTSI LFLVLNLT LTPDSNTTGT LSTSTASPH 540  
 RANSTSTERK LPEPESRGV IVAVIVCILV LAVLGAVLYF LYKKGKLPKR RSKQETILP 600  
 PSRKTELVEE VKSDKLPEEM GLLQSSGDK RAPGDQGEKY IDLRH

Seq ID NO: 554 DNA sequence  
Nucleic Acid Accession #: NM\_003183.3  
Coding sequence: 165..2639

65 1 11 21 31 41 51  
 TCGAGCTGG CGGTAGAATC TTCCAGTAG GCGGCGCGGG AGGGAAAAGA GGATTGAGGG 60  
 GCTAGGCGGG GCGGATCCCG TCTCCCGCG ATGTGAGCAG TTTCCGAAA CCCCGTCAGG 120  
 CGAAGGCTGC CCAGAGAGGT GGAGTCGGTA GCGGGCGCG GAACATGAG CAGTCTCTCC 180  
 TATTCTGAC CAGCGTGGTT CTTTCTGTC TGGCGCGCG ACCTCCGGAT GACCCGGGCT 240  
 70 TCGGCCCCA CCAGAGACTC GAGAAGCTTG ATTCTTTGCT CTCAGACTAC GATATTCTCT 300  
 CTTTATCTAA TATCCAGCAG CATTCGGTAA GAAAAAGAGA TCTACAGACT TCAACACATG 360  
 TAGAACAAC ACTAATCTTT TCAGCTTGA AAAGGCATT TAAATTATC CTGACATCAA 420  
 GTACTGAACG TTTTTCACAA AATTTCAGG TCGTGGTGGT GGATGGTAAA AACGAAAGCG 480  
 AGTACACTGC AAAATGGCAG GACTTCTTCA CTGGACCGT GGTGTGTGAG CCGTACTCTA 540  
 75 GGGTCTAGC CCACATAAGA GATGATGAT TTATAATCAG AATCAACACA GATGGGCGCG 600  
 AATATAACAT AGAGCCACTT TGGAGATTG TTAATGATC CAAAGACAAA AGAATGTTAG 660  
 TTTATAATC TGAAGATATC AAGAATGTT CACGTTTGA GTCTCCAAA GTGTGTGGTT 720  
 ATTTAAAGT GGATATGAA GAGTGTCTCC CAAAGGGTT AGTAGACAGA GAACCACTG 780  
 AAGAGCTTGT TCAATGAGT AAAAGAAGG CTGACCCAGA TCCCATGAG AACACGTGTA 840  
 80 AATTATTGTT GGTAGCAGAT CATCGCTCT ACAGATACAT TGACATCTAT CGGAACACTT 900  
 CAACTACAAA TTACTTAATA GAGCTAATTG ACAGAGTTGA TGACATCTAT CGGAACACTT 960  
 CATGGGATAA TGCAGGTTT AAAGGCTATG GAATACAGAT AGAGCAGATT CGCATTCTCA 1020  
 AGTCTCCACA AGAGGTAAAA CCTGGTGAAG AGCACTACAA CATGGCAAAA AGTTACCCAA 1080  
 ATGAAGAAAA GATGTCTGG GATGTGAAGA TGTGTCTAGA GCAATTAGC TTTGATATAG 1140  
 85 CTGAGGAAGC ATCTAAAGTT TGCTTGGCAC ACCTTTTCAC ATACCAAGAT TTTGATATAG 1200  
 GAACTCTGG ATTAGCTTAT GTTGGCTCTC CCAGAGCAAA CAGCCATGGA GGTGTTGTG 1260  
 CAAAGGCTTA TTATAGCCCA GTTGGGAGA AAAATATCTA TTTGAATAGT GGTGTTGACA 1320  
 GCACAAAGAA TTATGGTAAA ACCATCCTTA CAAAGGAAGC TGACCTGGTT ACAACTCATG 1380

	AATTGGGACA	TAATTTTGGG	GCAGAACATG	ATCCGGATGG	TCTAGCAGAA	TGTGCCCOGA	1440
	ATGAGGACCA	GGGAGGGA	TATGTCATGT	ATCCCATAGC	TGTGAGTGGC	GATCAGCAGA	1500
	ACATAAGAT	GTTCCTAAAC	TGCAGTAAAC	AATCAATCTA	TAAGACCAT	GAAGTAAGG	1560
5	CCAGAGAGTG	TTTTCAAGAA	CGCAGCAATA	AAGTTTGTGG	GAACCTGAGG	GTGGATGAAG	1620
	GAGAGAGTG	TGATCTCTGC	ATCATGTATC	TGAACAAACG	CACCTGTGTC	AACAGCGACT	1680
	GCACGTTGAA	GGAGGTGTGC	CAGTGCAGTG	ACAGGAACAG	TCCTTGCTGT	AAAAACTGTC	1740
	AGTTTGAGAC	TGCCCAAGAG	AAGTGCCAGG	AGGCGATTAA	TGCTACTTGC	AAAGCGTGT	1800
	CCTACTGCAC	AGGTAATAGC	AGTGAAGTCC	CGCCTCCAGG	AAATGCTGAA	AATGACACTG	1860
10	TTTGCTTGGG	TCTTGGCAAG	TGTAAGGATG	GGAAATGCAT	CCCTTTCTGC	GAGAGGGAAC	1920
	AGCAGCTGGA	GTCTCTGTCA	TGTAATGAAA	CTGCAAACTC	CTGCAAGGTG	TGCTGCAGGG	1980
	ACCTTTCTGG	CCGCTGTGTG	CCCTATGTGG	ATGCTGAACA	AAAGAACTTA	TTTTTGAGGA	2040
	AAGGAAAGCC	CTGTACAGTA	GGATTTTGTG	ACATGAATGG	CAATGTGTAG	AAACGAGTAC	2100
	AGGATGTAAT	TGAACGATTT	TGGGATTTC	TTGACCAGCT	GAGCATCAAT	ACTTTTGGAA	2160
15	AGTTTATAGC	AGACAACATC	GTGGGTCTG	TCCTGGTTT	CTCCTTGATA	TTTTGGATT	2220
	CTTTTACAGT	TCTTGTCCAT	TGTGTGGATA	AGAAATGGGA	TAAACAGTAT	GAATCTCTGT	2280
	CTCTGTTTCA	CCCAATGAA	GTGAAATGTC	TGAGCAGCAT	GATTCTGTCA	TCGTTTCCGA	2340
	TTATCAAAAC	CTTTCTGTGG	CCCCAGACTC	CAGGCCGCTC	GCAGCTGTCC	CCTGTGATCC	2400
	CTTCCGCGCC	AGCAGCTCCA	AAACTGGACC	ACCAGAGAA	GGACACCATC	CAGGAAGACC	2460
20	CCAGCACAGA	TCCCATATG	GACGAGGATG	GGTTTGAGAA	GGACCCCTTC	CCAAATAGCA	2520
	GCACAGCTGC	CAAGTCATTT	GAGGATCTCA	CGGACCATCC	GGTCCGCGAG	AGTGAAGAGG	2580
	CTGCTCTCTT	TAACTGTGAG	CGTCAGAAATC	GTGTTAACAG	CAAGAAACAA	GAGTGCTAAT	2640
	TTAGTTCTCA	GCCTCTGTGA	CTTAAGTGTG	CAAAATATTT	TTATAGATTT	GACCTACAAA	2700
	TCAATCACAG	CTTGATTTTT	GTGAAGACTG	GGAAGTGACT	TAGCAGATGC	TGGTCATGTG	2760
25	TTTGAACCTC	CTGCAGGTAA	ACAGTCTCTG	TGTGGTTTGG	CCCTTCTCCT	TTTGAAGAGG	2820
	TAAGGTGAAA	GTGAATCTAC	TTATTTTGAG	GCTTTCAGTG	TTTAGTTTTT	AAAATATCTT	2880
	TTGACCTGTG	CTGCAAAAGC	AGAAATACA	GCTGGATTGG	GTATGAATA	TTTAGTTTTT	2940
	TGTAATTTAA	TCTTTTATAT	TGTAACAGC	ACTGACTAGG	GAAATGATCA	GTTTTTTTTT	3000
	ATACACTGTA	ATGAACCGCT	GAATATGAAG	CAATTGGCAT	TTATTTGTGA	GAAAGTGGAA	3060
30	ATAGTTTTTT	TTTTTTTTTT	TTTTTTTTTG	CTTCAACTAA	AAACAAAGGA	GATAAATTTA	3120
	GTATACATGT	TATCTAAAT	GTGGGCTAT	TTCTAGTTAT	TACCCAGAGT	TTTTATGTAG	3180
	CAGGGAATAT	ATATATCTAA	ATTTAGAAAT	CAATTGGGTT	AATATGGCTC	TTTCAATTTT	3240
	TAAGACTAAT	GCTCAGAAAC	TAACCACTAC	CTTACAGTGA	GGGCTATACA	TGGTAGCCAG	3300
	TTGAATTTAT	GGAATCTACC	AACTGTTTAG	GGCCCTGATT	TGCTGGGCAG	TTTTTCTGTA	3360
35	TTTTATAAGT	ATCTTCATGT	ATCCCTGTTA	CTGATAGGGA	TACATGTCTT	AGAAAATCTA	3420
	CTATTGGCTG	GGAGTGGTGG	CTCATGCCCT	TAATCCAGC	ACTTGGAGAG	GCTGAGGTTG	3480
	CGCCACTACA	CTCAGCCTTG	GGTGACAGAG	TGAGATCTGC	CTC		

Seq ID NO: 555 Protein sequence  
Protein Accession #: NP\_003174.2

40	1	11	21	31	41	51	
	MRQSLLFLTS	VVPFVLAPRP	PDDPGPGPHQ	RLEKLDLSLS	DYDILSLNSI	QQHSVRKRDL	60
45	QTSTHVELLL	TPSALKRHF	LYLTSSTERF	SONFRVUVVD	GKNESEYIAK	WQDFPTGHVV	120
	GEPSDRVLAH	IRDDDDVIIR	NTDGAENIE	PLWRFVNDTK	DKRMLVYKSE	DIKNVSRILQS	180
	PKVCGYLVVD	NEELLKPLGV	DREPPPELVH	RVKRRADPDP	MKNTCKLLVV	ADHRFYRYMG	240
	RGEESTTNY	LIELIDRVDD	IYRNTSWDNA	GFGYGIQIE	QIRILKSPQE	VKPEKHYNM	300
	AKSYPMNEEK	ANDVKMLLEQ	FSFDIAEES	KVCLAHFLTY	QDFDMGTGL	AYVGSPPRNS	360
50	HGVCPCPKAY	SPVGKKNYIL	NSGLTSTKNY	GKTLTREAD	LVTTHLGHN	FGAEDPDGL	420
	AECAPNEDQG	GKVMYPIAV	SGDHEHNMKF	SNCSKQSIYK	TIESKAQECF	QERSNKKVGN	480
	SRVDEGEBCD	PGIMLYLNDT	CNSDCTLKE	GVQCSDRNSP	CKNCQFETA	QKKQCEAINA	540
	TKGVSYCTG	NSSECPPPGN	AENDTVCLDL	GKCKDGKICP	PCEREQQLS	CACNETDNCS	600
	KVCCRLDSGR	CVFYVDAAEQ	NLFRLKRGKPC	TVGFCDMNGK	CEKRVQDVIE	RFWDFIDOLS	660
55	INTFGFLAD	NIVGSLVFS	LIFWIPFSIL	VHCVDKLKD	QYESLSLFEP	SNVEMLSMD	720
	SASVRIKFP	PAPQTPGRLL	PAPVIPSAPA	APKLDEQRMD	TIQEDPSTDS	HMDDEGFED	780
	PFPSNSTAAK	SFEDLTDHPV	ARSEKAASF	LQRQNRVNSK	ETEC		

Seq ID NO: 556 DNA sequence  
Nucleic Acid Accession #: NM\_021832.1  
Coding sequence: 164..2248

60	1	11	21	31	41	51	
	TCGAGCCTGG	CGGTAGAATC	TTCCAGTAG	GCGGCGCGGG	AGGAAAAGAG	GATTGAGGGG	60
65	CTAGGCGGGG	CGGATCCCGT	CCTCCCGGGA	TGTGAGCAGT	TTTCCGAAAC	CCCGTCAGGC	120
	GAAGGCTGCC	CAGAGAGGTG	GAGTCGGTAG	CGGCGCGGGG	AACATGAGGC	AGTCTCTCCT	180
	ATTCTGACC	AGCGTGGTTC	CTTTGTGTCT	GGCGCCGCGA	CCTCCGGATG	ACCCGGGCTT	240
	CGGCCCCAC	CAGAGACTCG	AGAAGCTTGA	TTCTTTGTCT	TCAGACTACG	ATATTCTCTC	300
70	TTTATCTAAT	ATCCAGCAGC	ATTCCGTAAG	AAAAGAGAT	CTACAGACTT	CAACACATGT	360
	AGAAACACTA	CTAACTTTT	CAGCTTTGAA	AAGGCATTTT	AAATTATACC	TGACATCAAG	420
	TACTGAACGT	TTTTCAAAA	ATTTCAAGGT	CGTGTGGTGT	GATGTGAAAA	ACGAAAGCGA	480
	GTACACTGTA	AAATGGCAGG	ACTTCTTCAC	TGGACACGTG	GTTGGTGAAC	CTGACTCTAG	540
	GGTTCTAGCC	CACATAAGAG	ATGATGATGT	TATAATCAGA	ATCAACACAG	ATGGGGCGGA	600
75	ATATAACATA	GAGCCACTTT	GGAGATTGT	TAATGATACC	AAAGACAAAA	GAATGTTAGT	660
	TTATAAATCT	GAGATATCA	AGAATGTTTC	ACGTTTGAGT	TCTCCAAAAG	TGTGTGGTTA	720
	TTTAAAGGTG	GATAATGAAG	AGTTGCTCCC	AAAAGGGTTA	GTAGACAGAG	AACCACTGTA	780
	AGAGCTTGT	CATCGAGTGA	AAAGAAGAGC	TGACCCAGAT	CCCATGAAGA	ACACGTGTAA	840
	ATTATTGGTG	GTAGCAGATC	ATCGCTTCTA	CAGATACATG	GGCAGAGGGG	AGAGAGTAC	900
80	AACACAAAT	TACTTAATAG	AGCTAATTGA	CAGAGTTGAT	GACATCTATC	GGAAACACTTC	960
	ATGGGATAAT	GACGGTTTGA	AAGGCTATGG	AATACAGATA	GAGCAGATTG	GCAATCTCAA	1020
	GTCTCCACAA	GAGGTAAAAA	CTGGTAAAAA	GCACTACAAC	ATGGCAAAAA	GTATCCCAAA	1080
	TGAAGAAAAG	GATGCTTGGG	ATGTGAAGAT	GTGCTAGAG	CAATTAGCT	TTGATATAGC	1140
	TGAGGAAGCA	TCTAAGATT	GCTTGGCACA	CCTTTTCACA	TACCAAGATT	TTGATATGGG	1200
85	AACCTCTGGA	TTAGCTTATG	TGGCTCTCC	CAGAGCAAA	AGCCATGGAG	GTGTTTGTCC	1260
	AAAGGCTTAT	TATAGCCAGC	TTGGGAAGAA	AAATATCTAT	TTGAATAGTG	GTGTTGACAG	1320
	CACAAAGAA	TATGTTAAAA	CCATCCTTAC	AAAGGAAGCT	GACCTGGTTA	CAACTCATGA	1380
	ATTGGGACAT	AATTTTGGAG	CAGAACATGA	TCCGGATGGT	CTAGCAGAA	GTGCCCCGAA	1440

5  
 10  
 15  
 20  
 25  
 30  
 35

TGAGGACCAG GGAGGGAAAT ATGTCATGTA TCCCATAGCT GTGAGTGGCG ATCACGAGAA 1500  
 CAATAAGATG TTTTCAAACT GCAGTAACA ATCAATCTAT AAGACCATTT AAAGTAAGGC 1560  
 CCAGGAGTGT TTTCAAGAAC GCAGCAATAA AGTTTGTGGG AACTCGAGGG TGGATGAAGG 1620  
 AGAAGAGTGT GATCTGGCA TCATGTATCT GAACAACGAC ACCTGCTGCA ACAGOGACTG 1680  
 CACGTTGAAG GAAGGTGTCC AGTGCACTGA CAGGAACAGT CCTTGTCTGA AAAACTGTCA 1740  
 GTTTGAGACT GCCCAGAAGA AGTGCCAGGA GCGGATTAAT GCTACTTGCA AAGGCGGTGC 1800  
 CTACTGCACA GGTAAATAGCA GTGAGTGCCC GCCTCCAGGA AATGCTGAAG ATGCACTGT 1860  
 TTGCTTGGAT CTGTCGAAGT GTAAGGATGG GAAATGCATC CCTTCTGCG AGAGGGAAAC 1920  
 GCAGCTGGAG TCCTGTGCAT GTAATGAAAC TGACAACTCC TGCAAGGTGT GCTGCAGGGA 1980  
 CCTTTCCGGC CGCTGTGTGC CCTATGTCGA TGCTGAACAA AAGAACTTAT TTTTGAGGAA 2040  
 AGGAAAGCCC TGTACAGTAG GATTTTGTGA CATGAATGCG AAATGTGAGA AACGAGTACA 2100  
 GGATGTAATT GAACGATTTT GGGATTTTCT TGACCACTG AGCATCAATA CTTTGGGAAA 2160  
 GTTTTTCAGCA GACAACATCG TTGGGTCTGT CTGTGTTTCT TCCTTGATAT TTTGGATTCC 2220  
 TTTTTCAGATT CTGTGCAATT GTGTGTAACG TCGAAATGCT GAGCAGCATG GATTCTGCAT 2280  
 CGGTTCCGAT TATCAAAACC TTCTCTGGC CCCAGACTCC AGGCCGCTG CAGCCTGCCC 2340  
 CTGTGATCCC TTGCGCGCCA GCAGCTCCAA AACTGGACCA CCAGAGAATG GACACCATCC 2400  
 AGGAAGAGCC CAGCACAGAC TCACATATGG ACAGGATGG GTTTCAGAAG GACCCTTCC 2460  
 CAAATAGCAG CACAGCTGCC AAGTCATTGG AGGATCTCAC GGACCATCCG GTCACCAGAA 2520  
 GTGAAAGGCG TGCTCTCTTT AAACCTGCAGC GTCAGAATCG TGTTCAGAGC AAAGAAACAG 2580  
 AGTGCTAATT TAGTCTCAG CTCTCTGAC TTAAGTGTGC AAAATATTTT TATAGATTG 2640  
 ACCTACAATC AATCAGCTG TATATTTTGT GAAGACTGGG AAGTGACTTA GCAGATGCTG 2700  
 GTCATGTGTT TGAACCTTCT GCAGGTAAAC AGTTCTTGTG TGGTTTGGCC CTCTCTCTT 2760  
 TGAAGAGTGA AGGTGAAGGT GAATCTAGCT TATTTTGAGG CTTTCAGGTT TTATGTTTTA 2820  
 AAATATCTTT TGACCTGTGG TGCAAAAGCA GAAATACAG CTGGATTGGG TTATGAGTAT 2880  
 TTAAGTTTTT GTAATTAAT CTTTTATATT GATAACAGCC ACTGACTAGG GAAATGATCA 2940  
 GTTTTTTTTT ATACACTGTA ATGAACCGCT GAATATGAAG CATTGGCAT TTATTTGTGA 3000  
 GAAAAGTGGG ATAGTTTTTT TTTTTTTTTT TTTTTTTTGC CTTCACCTAA AAACAAGGA 3060  
 GATAAATTTA GTATACATG TATCTAAAT GTGGGTCTAT TTCTAGTTAT TACCCAGAT 3120  
 TTTTATGTAG CAGGAAAAT ATATATCTAA ATTTAGAAAT CATTGGGTTT AATATGGCTC 3180  
 TTTTATCTG TAAGACTAAT GCTCAGAAC TAACCACTAC CTTACAGTGA GGGCTATACA 3240  
 TGGTAGCCAG TTGAATTAT GGAATCTACC AACTGTTTAG GGCCCTGATT TGCTGGGCG 3300  
 TTTTCTGTA TTTTATAAGT ATCTTCATGT ATCCCTGTTA CTGATAGGGA TACATGTCTT 3360  
 AGAAAATTCA CTATTGGCTG GAGTGGTGG CTCATGCCCT TAATCCAGC ACTTGAGAG 3420  
 3421 GCTGAGGTTG CCCCCTACA CTCAGCCTG GGTGACAGG TGAGATCTGC CTC

Seq ID NO: 557 Protein sequence  
 Protein Accession #: NP\_068604.1

40  
 45  
 50

1 11 21 31 41 51  
 MRQSLFLTS VVFPVLAPRP PDDPGFGPEQ RLEKLDLSLS DYDILSLNI QHRSVRKRD 60  
 QTSTHVELLL TFSALKRHF LYLTSSTERF SQNFVYVVD GKNSEHYTVK WQDFPTGHVV 120  
 GEPDSRVLAH IRDDVDIIRI NTGAEYNIE PLNRFPVNDK DKRLVYKSE DIKNVSLQS 180  
 PKVGYLKV DREPPPEELVE RVKRRADPDF MKRTCKLLV ADHRFYRYMG 240  
 RGEESTTNY LIELIDRVDD IYRNTSWDNA GFKGYGQIE QIRILKSPQE VKPGEKHYNM 300  
 AKSPNEEDK AWDVQMLLEQ FSPDIAEEAS KVCLAHLPY QDFDMGTGL AYVGSPPRANS 360  
 HGVGCKATY SPVQKNIYL NSGLSTKNY GKTILTKBAD LVTTHLGHN FGAEDPDGL 420  
 AECAPNEDQ GKYVMYPIAV SGDHENKMF SNCSKQSIYK TIESKAQECF QERSNKKVGN 480  
 SRVDEGEEDC PGIMLYNNDT CCNSDCTLKE GVQCSDRNSP CCKNQCFETA QKKQCEAINA 540  
 TKGVSYCTG NSSECTPPPP AEDDTVCLDL GKCHDGKCP FCEREQQLBS CACNETDNCS 600  
 KVCCRDLSGR CVPYDAEQK NLFLRKGP KCVRVQDVIE RFWDFIDQLS 660  
 INTFGKFLAD NIVGSVLVFS LIPWIPFSL VHCV

Seq ID NO: 558 DNA sequence  
 Nucleic Acid Accession #: NM\_004994.1  
 Coding sequence: 20..2143

60  
 65  
 70  
 75  
 80  
 85

1 11 21 31 41 51  
 AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCTCG GTCTGTGTGC TCTGTGTGCT 60  
 GGGCTGCTGC TTTGCTGCC CCAGACAGCG CCAGTCCACC CTGTGTCTCT TCCCTGGAGA 120  
 CCTGAGAAC AATCTCACCG ACAGGACAGT GGCAGAGGAA TACCTGTACC GCTATGGTTA 180  
 CACTCGGGT GCAGAGATGC GTGGAGATGC GAAATCTCTG GGGCTGCGC TGCTGTCTCT 240  
 CCAGAAGCAA CTGTCCCTGC CCGAGACCGG TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300  
 GCGAACCCCA CGGTCCGGGG TCCAGACCT GGGCAGATTC CAAACCTTTG AGGGCGACCT 360  
 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAC TACTCGGAAG ACTTGGCCGG 420  
 GGCCTGATT GACGACCCCT TTGCCCGCGC CTTCGCACTG TGGAGCGCGG TGACGCCGCT 480  
 CACCTTCACT CGCGTGACA GCGGGGACGC AGACATCGTC ATCCAGTTTG GTGTGCGGGA 540  
 GCAGGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACAGC CCTTCTCTCC 600  
 TGGCCCCGGC ATTCAGGGAG ACGCCATTIT CGACGATGAC GAGTTGTGGT CCTTGGGCAA 660  
 GGGCGTCTGT GTTCCAATC GGTTTGGAAA CGCAGATGGC GCGGCTGCG ACTTCCCTTT 720  
 CATCTTCGAG GGCCTCTCT ACTCTGCGTG CACCACCGAC GGTGCTGCG ACAGCTTGCC 780  
 CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTTGGCTTCT GCCCCAGCGA 840  
 GAGACTCTAC ACCCGGGAGC GCAATGTCTG TGGGAAACCC TGCCAGTTTC CATTCTCTT 900  
 CCNAGGCCAA TCCCTACTCG CCTGCACCAC GACAGGTGCG TCCGACGGCT ACCGCTGGTG 960  
 CGCCACCACC GCCAACTACG ACCGGGACAA GCTCTTGGG TTCTGCCCGA CCGAGCTGA 1020  
 CTGACGGGTG ATGGGGGGCA ACTCGGGGGG GAGCTGTGCG GTCTTCCCT TCACTTTCTC 1080  
 GGGTAAGGAG TACTCGACCT GTACCAAGCA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1140  
 TACCACCTCG AACTTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200  
 TTTGTTCTCT GTGGCGGCGC ATGAGTTCCG CCACGCGCTG GGCTTAGATC ATTCTCAGT 1260  
 GCGGAGGCG CTCTGTATCC CTATGTACCG CTTCAGTGG GGGCCCCCTT TGCATAAGGA 1320  
 CGAGTGAAT GGCATCGGCG ACCTCTATGG TCCTGCGCTT GAACCTGAGC CACGGCCTCC 1380  
 AACCAACACC ACACCGGCTC CCGGACGCTC TGCCCAACCG GACCCCTCAC 1440  
 TGTCCACCCC TCAGAGCGCC CACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC 1500  
 AGGTCCCCCC ACTGCTGGCG CTTCTACGGC CACTACTGTG CCTTTGAGTC CGGTGGACGA 1560  
 TGCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAACACAG TGTATTGTT 1620  
 CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGG AGCGGGCCG AGGGCCCCCT 1680

CCTATCGGCC GACAACTGGC CGCGCTGCC CGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740  
 GCTCTCCAAG AAGCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC 1800  
 GGTGCTGGCC CCGAGGCGTC TGGACAAGCT GGGCTGGGA GCCGACGTGG CCCAGGTGAC 1860  
 CGGGGCCCTC CGGAGTGGCA GGGGGAAGAT GCTGTGTTC AGCGGGCGGC GCCTCTGGAG 1920  
 GTTCTAGCTG AAGCGCGAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980  
 CCGCGGGGTG CCTTTGGACA CGCAGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG 2040  
 CCAGGACCGC TTCTACTGGC GCGTGAGTTC CCGGAGTGG TTGAACCCAG TGGACCAAGT 2100  
 GGGCTACGTG ACCTATGACA TCCTGAGTGG CCCTGAGGAC TAGGGCTCCC GTCTGCTTTT 2160  
 GCAGTGCATG GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCCGGATA 2220  
 CAAACTGGTA TTCTGTCTCG GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC 2280  
 TCACCTTTGT TTTTGTGTGG AGTGTCTTCTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 559 Protein sequence  
 Protein Accession #: NP\_004985.1

1 11 21 31 41 51  
 MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRLNLT DRQLAEVLY RYGYTRVAEM 60  
 RGSKSLSGPA LLLLQKQLSL PETGELDSAT LKAMRTPRCG VPDLGRPQTF EGDLEWHHN 120  
 ITYWIQNYSE DLPRVIDDA FARAPALWSA VTPLTFTRVY SRDADIVIQF GVAEHGQYYP 180  
 PDGKDLALLAH APFPGPIQGG DAHPDDDELW SLGKGVVVPF RFGNADGAAC HFFPIFEGRS 240  
 YSACTTDGRS DGLPWCSTTA NYDTDDRFGF CPSELYTRD GNADGKPCQF PFIFQGSYS 300  
 ACTTDGRSDG YRWCAATANY DRDKLFGPCP TRADSTVMGG NSAGELCVFP FTFLGKEYST 360  
 CTSEGRDGR LWCATTNFD SDKKWGFCPD QGYSLFLVAA HEFHALGLD HSSVPEALMY 420  
 PMYRFTGEPF LHKDDVNGIR HLYGPRPEPE PRPPTTTTQ PTAPPTVCPT GPPTVHPSE 480  
 PTAGPTGPPS AGPTGPTTAG PSTATTPLS PVDDACNVNI FDAIAEIGNQ LYLFDKQKYN 540  
 RFSBGRSRP QGPFLLADKW PALPRKLDV FEEPLSKFLP FSGRQVWVY TQASVLGPRR 600  
 LDKLGLGADV AQTGALRSB RGHMLFSGR RLWRFVKAQ MVDPRSASEV DRMFPGVPLD 660  
 THDVQYREK AYFCQDRFYW RVSSRSRLNQ VDQGVVITYD ILQCPED

Seq ID NO: 560 DNA sequence  
 Nucleic Acid Accession #: NM\_000213.1  
 Coding sequence: 127..5385

35 1 11 21 31 41 51  
 CGCCCGCGCG CTGCAGCCCC ATCTCTAGC GGCAGCCAG GCGCGGAGG AGCGAGTCCG 60  
 CCCCAGGTA GGTCCAGGAC GGGCGCACAG CAGCAGCCGA GCTGGCCGG GAGAGGGAGG 120  
 AAGAGGATGG CAGGGCCACG CCCACGCCCA TGGGCCAGGC TGCTCTGGC AGCCTTGATC 180  
 AGCGTCAGCC TCTCTGGGAC CTTGGCAAC CGCTGCAAGA AGGCCCCAGT GAAGAGCTGC 240  
 ACCGAGTGTG TCCGTGTGGA TAAGGACTGC GCCTACTGCA CAGACGAGAT GTTCAGGGAC 300  
 CGGCGCTGCA ACACCCAGCG GGAGCTGCTG CCGCGCGGCT GCCACGGGGA GAGCATCGTG 360  
 GTCATGGAGA CAGGCTTCCA AATCACAGAG GAGACCCAGA TTGACACAC CCGTGGCGCG 420  
 AGCCAGATGT CCCCCAAGG CCGCGGGTGC CGTCTGCGC CCGGTGAGGA GCGGCATTTT 480  
 GAGCTGGAGG TGTTTGAAGC ACTGGAGAGC CCGGTGAGC TGTACATCCT CATGGACTTC 540  
 TCCAACCTCA TGTCCGATGA TCTGGACAA CTCAAGAAGA TGGGGCAGAA CCGTGGCTCG 600  
 GCTCTGAGCC AGCTACACAG CGACTACACT ATTGGATTGG GCAAGTTTGT GGACAAAGTC 660  
 AGCGTCCCGC AGACGACACT GAGGCTGAG AAGCTGAAG AGCCCTGGCC CAACAGTGAC 720  
 CCCCCCTTCT CTTTCAAGAA CGTCATCAGC CTGACAGAAG ATGTGGATGA GTTCCGGAAT 780  
 AACTGCGAGG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CTGAGGGGG CTTCGATGCC 840  
 ATCCTGCAGA CAGCTGTGTG CACGAGGAGC ATTGGCTGGC GCGCCGACAG CACCCACCTG 900  
 CTGGTCTTCT CCACCGAGTC AGCCTTCCAC TATGAGGCTG ATGGGCGCAA CGTGTGGCT 960  
 GGCATCATGA GCGCAACGGA TGAACGGTGC CACCTGGACA CCACGGGCAC CTACACCCAG 1020  
 TACAGACAC AGGACTACCC GTGCGTCCCC ACCCTGGTGC GCCTGCTGCG CAAGCACAA 1080  
 ATCATCCCCA TCTTTGCTGT CACCAACTAC TCCTATAGCT ACTACGAGAA GCTTCACACC 1140  
 TATTTCCCTG TCTCTCACT GGGGGTGTG CAGGAGSACT CGTCCACAT CGTGGAGCTG 1200  
 CTGGAGGAGG CTTTCAATCG GATCCGCTCC AACCTGGACA TCGGGCCCT AGACAGCCCC 1260  
 CGAGGCTTCT GAGACAGAGT CACCTCCAAG ATGTTCCAGA AGACGAGGAC TGGGTCTTTT 1320  
 CACATCCGGC GGGGGGAAGT GGGTATATAC CAGGTGCAGC TGGGGCCCT TGAGCACGTG 1380  
 GATGGAGGAC AGGTGTGCCA GCTGCGGAG GACCAGAAG GCAACATCCA TCTGAAACCT 1440  
 TCCTTCTCCG ACGGCTCAA GATGGAGCG GGCATCATCT GTGATGTGTG CACCTGCCAG 1500  
 CTGCAAAAAG AGGTGCGGTC AGCTCGCTGC AGCTTCAAG GAGACTTCGT GTGCGGACAG 1560  
 TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCG CTCTCTGAGT 1620  
 GACATTCAGC CCGCTCTGCG GAGGGGCGAG GACAAAGCGT GCTCCGGCG TGGGAGTGC 1680  
 CAGTGGCGGC ACTGTGTGTG CTACGGGAG GGCCTACG AGGGTCAGTT CTGCGAGTAT 1740  
 GACAACTTCC AGTGTCCCG CACTTCCGGG TTCTCTGCA ATGACGAGG ACCTGCTCC 1800  
 ATGGCCAGT GTGTGTGTGA GCCTGGTGG ACAGGCCCAA GCTGTGACTG TCCCTCAGC 1860  
 AATGCCACCT GACATGACAG CAATGGGGGC ATCTGTAATG GACGTGGCCA CTGTGAGTGT 1920  
 GCGCGCTGCC ACTGCCACCA GAGTGCCTC TACACGACA CCATCTGCGA GATCAACTAC 1980  
 TCGGCGATCC ACCCGGGCTC CTGCGAGGAC CTACGCTCT GGTGCAAGT CCAGGCGTGG 2040  
 GGCACCGCGC AGAAGAAAGG GCGCACGTGT GAGGAATGCA ACTTCAAGGT CAAGATGGTG 2100  
 GACGAGCTTA AGAGAGCCGA GGAGTGGTG GTGCGCTGCT CTTCCGGGA CGAGGATGAC 2160  
 GACTGCACCT ACAGCTACAC CATGGAAGT GACGGGCGCC CTGGGCCCAA CAGCACTGTC 2220  
 CTGGTGACCA AGAAGAAAGA CTGCCCTCG GGTCTCTCT GTTGGCTCAT CCCCCTGCTC 2280  
 TCTCTCTTCC TGCGCTCTCT GGCCTTGCTA CTGCTGCTAT GCTGGAAGTA CTGTGCTGTC 2340  
 TGCAAGGCTT GCCTGGCACT TCTCCGTGTC TGCAACGAG GTCAATGGT GGGCTTTAAG 2400  
 GAAGACCACT ACATGCTGCG GGAGAACCTG ATGGCTCTG ACCACTTGA CACGCCCATG 2460  
 CTGCGCAGCG GGAACCTCAA GGGCGGTGAC GTGGTCCGCT GGAAGGTAC CAACAACATG 2520  
 CAGCGGCTCG GCTTTGCCAC TCATGCGGCC AGCATCAACC CCACAGAGCT GGTGCCCTAC 2580  
 GGGCTGTCTT TGCGCTCTCT CCGCTTTGTC ACCGAGAACC TGCTGAAGCC TGACACTGCG 2640  
 GAGTGGCGCC AGCTGCGCCA GGAGGTGGAG GAGAACCCTA ACGAGTCTA CAGGAGATC 2700  
 TCGGTGTAC ACAAGCTCCA GCAGACCAAG TTCCGGCAGC AGCCCAATGC CGGGAATAAG 2760  
 CAAGACCACA CCATTGTGGA CACAGTGTG ATGGCGCCCC GCTCGGCCAA GCGGCGCTG 2820  
 CTGAAGCTTA CAGAGAAAGC GGTGGAAAC AGGGCTTTC ACAGCTCAA GGTGGCCCC 2880  
 GGCTACTACA CCGTCACTGC AGACCAAGAC GCGCGGGGCA TGTGGAGTT CCAGGAGGCG 2940  
 GTGGAGCTGG TGGAGTACG GGTGCCCTC TTTATCCGCG CTGAGGATGA CAGCAGAGAG 3000  
 CAGCTGCTGG TGGAGGCAT CAGCTGCCC GCAGGCAGT CCACCTCGG CCGCGGCTG 3060

5  
10  
15  
20  
25  
30  
35  
40  
45

GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGACGTGG TGTCTTTGA GCAGCCTGAG 3120  
TTCTCGGTCA GCCCGGGGGA CCAGGTGGCC GCATCCCTCG TCATCCGGCG TGTCTTGGAC 3180  
GGCGGGAGT CCCAGGTCTC CTACCGCACA CAGGATGGCA CGCGCGAGGG CAACCGGGAC 3240  
TACATCCCGG TGGAGGGTGA GCTGCTGTT CAGCCTGGGG AGGCTCGGAA AGAGCTGCAG 3300  
GTGAAGCTCC TGGAGCTGCA AGAAGTTGAC TCCCTCCTGC GGGGCGGCCA GGTCCGCGGT 3360  
TTCACGTCC AGCTCAGCAA CCTAAGITT GGGGCCACCC TGGCGCAGCC CCACTCCACC 3420  
ACCATCATCA TCAGGGAGCC AGATGAAGTG GACCGGAGCT TCACGAGTCA GATGTTGTCA 3480  
TCACAGCCAC CCCCTCAGCG GACCTGGGCG GCCCGGAGCA ACCCAATGC TAAGGCCGCT 3540  
GGGTCCAGGA AGATCCATT CAACTGGCTG CCCCCTTCTG GCAAGCCAAT GGGGTACAGG 3600  
GTAAAGTACT GGAATCAGGG TGAATCCGAA TCCGAAGCCC ACCTGCTCGA CAGCAAGGTG 3660  
CCCTCAGTGG AGCTCACCAG CCTGTACCCG TATTGCGACT ATGAGATGAA GGTGTGCGCC 3720  
TACGGGGCTC AGGCGAGGG ACCCTACAGC TCCCTGGTGT CTGCGCGAC CCACCGAGAA 3780  
GTGCCAGCG AGCCAGGGCG TCTGGCCTTC AATGTCGTCT CCTCCACGGT GACCCAGCTG 3840  
AGCTGGGCTG AGCCGGCTGA GACCAACGGT GAGATCACAG CCTACGAGGT CTGCTATGGC 3900  
CTGGTCAAG ATGCAACCCG ACCTATTGGG CCAATGAAGA AGTGCTGGT TGACAACCTC 3960  
AAGAAGCGGA TGGTCTTAT TGAGAACCTT CGGAGTCCC AGCCCTACCG CTACACGGTG 4020  
AAGGCGGCA AGGCGGCGG CTGGGGGCTC GAGCGGGAGG CCATCATCAA CCTGGCCACC 4080  
CAGCCCAAGA GGCCCATGTC CATCCCCATC ATCCCTGACA TCCCTATCGT GGACGCCAG 4140  
AGCGGGAGG ACTAGACAG CTTCTTATG TACAGCGATG AGTTCACG CTCTCCATCG 4200  
GGCAGCCAGA GGCCAGCGGT CTCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAC 4260  
TTTGCTTCC CGGCGAGCAG CAACTCCCTG CACAGGATGA CCACGACCA TGTGCTGCC 4320  
TATGGCACC ACCTGAGCCC ACAGTGGCC CACCGCGTGC TAAGCACATC CTCCACCTC 4380  
ACAGGGAGT ACAACTCACT GACCGCTCA GAACACTCAC ACTCGACCAC ACTGCGAGG 4440  
GACTACTCCA CCTCAGCTC CGTCTCCTC CAGACTCTC GCCTGACTGC TGGTGTGCC 4500  
GACAGCCCA CCGCCTGGT GTTCTCTGCC CTGGGGCCCA CATCTCTCAG AGTGAGCTGG 4560  
CAGGAGCGCG TCGTGGAGG GCGCTGCGAG GGCTACAGT TGGAGTACCA GCTGCTGAAC 4620  
GGCGGTGAGC TGATCGGCT CAACATCCCC AACCTGCCCC AGACCTCGGT GGTGTGGAA 4680  
GACCTCTGC CCAACCACTC CTACGTGTT CCGTGGCGGG CCGAGGCCA GGAAGCTGG 4740  
GGCCGAGAGC GTGAGGGTGT CATCAACATT GAATCCCAGG TGCACCCGCA GAGCCCACTG 4800  
TGTCCCTGC CAGGCTCCGC CTCTACTTTG AGCACTCCCA GTGCCCCAGG CCCGCTGGTG 4860  
TTCACTGCC TCGAGCCAGA CTGCTGCGAG CTGAGCTGGG AGCGGCCACG GAGGCCCAAT 4920  
GGGATATCG TCGGCTACT GTGACCTGT GAGATGGCCC AAGGAGGAGG GCCAGCCACC 4980  
GCATTCCGGG TGAATGAGGA CAGCCCGAG AGCCGCTGA CCGTGGCGGG CCTGAGCAG 5040  
AACGTGCCCT ACAAGTTCAA GGTGAGGCC AGGACCACTG AGGCGCTCGG GCCAGAGCGC 5100  
GAGGGCATCA TCACCATAGA TCCCAAGGAT GGAGGACCTC TCCCGAGCT GGGCAGCGT 5160  
GCCGGGCTCT TCCAGCACCC GCTGCAAGC GAGTACAGCA GCATCACACC CACCCACACC 5220  
AGCGCCACCG AGCCCTTCTC AGTGAATGGG CCGACCTGGG GGGCCCGACA CCTGAGGCA 5280  
GGCGGCTCCC TCACCCGCA TGTGACCCAG GAGTTTGTGA GCGGACACT GACCCACAGC 5340  
GGAACCCCTA GCACCCACAT GGACCAACAG TTCTTCCAAA CTGACCGCA CCTTGCCTCA 5400  
CCCCCGCAT GTCCCACTAG CGGTCTCTCC GACTCTCTC CCGAGGCTC CTCAGTACT 5460  
CCATCTTGC ATCCCTGGG GCGGACCCCA CCGCATGCA CAGAGCAGG GCTAGGTGTC 5520  
TCTGGGAGG CATGAAGGG GCAAGTCCG TCCTCTGTGG GCCCAAACT ATTTGTAACC 5580  
AAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGGTTTTC 5640  
TACTG

Seq ID NO: 561 Protein sequence  
Protein Accession #: NP\_000204.1

50  
55  
60  
65  
70  
75  
80

1 11 21 31 41 51  
MAGPRSPWA RLLLAALISV SLSTLANRC KKAPVKSCTE CVRVKDCAY CTDEMFRDRR 60  
CNTQAEALAA GCGRESIVVM ESSFQITEET QIDTLRRSQ MSPQGLRVL RPGEERHPEL 120  
EVFEPLSEPV DLYILMDFSN SMSDDLNLK KMGQNLARVL SQLTSDYITG PGKFDVKVSV 180  
PQTDMPERKL KEPWNSDEP FSPKIVISLT EDVDEFNKL QGERISGNLD APEGGFDAIL 240  
QTAVCTRDIG WRPDSTHLV FSTESAFHYE ADGANVLAGE MSRNDERCHL DTGTGTYTQR 300  
TQDYPSPVPTL VRLLAGHNI PIFAVTNSY SYEKLTHTYF FVSSLGVLQE DSSNIVELLE 360  
EAFNRIRSNL DIRALDSPRG LRTEVTSKMF QKTRTGSPHI RRGEVGIYQV QLRALSHVDG 420  
THVCQLPEQD GKNHILKPSF SDGLKMDAGI ICDVCTCELO KEVRSARCSF NGDFVCGQCV 480  
CSBGMGGQTC NCSTGSLSDI QPCLREGEDK PCSGRGECQC GHCVCYGEGR YBGFCEYDN 540  
FQCPRTSGLP CNDRGRCMSG QCVCEPGWTG PSCDCPLENA TCIDSNGGIC NGRGHCCEGR 600  
CHCHQQSLYT DTICEINYS A IHPGLCEDLR SCVQCQAWT GEKKGRCTCE CNPKVMVDE 660  
LKRAEEVVVR CSFRDEDDC TSYSTMEDG APGPNSTVLV HKKDCPPGS FWWLIPLLL 720  
LLPLALLLLL LCWKYCACC ACIALLPCCN RGHMVGFKED HYMLRENLM SDHLDTPLMR 780  
SENLKGRDVV RWEVTNNMQR PGFATHAASI NPTELVPYGL SLRLARLCTE NLLKPDTRC 840  
AQLRQVEVEN LNEVYRQISG VHKLQTKER QPNAGKKQD HTIVDTVLM PRSAKPALLK 900  
LTEKQVEQRA FHDLVKAPGY YTLTADQDAR GMVEPQEGVE LVDVRVPLFI RPEDDDEKQL 960  
LVEAIDVPAG TATLGRRLVN ITIIEQARD VVSFEQPEFS VSRGDQVARI FVIRRVLDGG 1020  
KSQVSYRTQD GTAGGNRDI PVGEGLLPQ GEAWKELQVK LLELQEVDSL LRGRQVRPF 1080  
VQLSNPKFGA HLGQPHSTTI IIRDPDELDR SFTSOMLSSQ PPPHGLDGLAP QNPNAKAAGS 1140  
RKIHFNWLPF SGKPMGYRVK YWIQDSESE AHLDSKVPS VELTNLYPYC DYEMKVCAYG 1200  
AQGEQPYSSL VSCRTHQEV SEPRGLAFNV VSSTVTQLSW AEPATNGEI TAYEVCYGLV 1260  
NDDNRPIGPM KKVLDNPKN RMLLENLRE SQPYRYTVKA RAGAGWGER EAINLATQP 1320  
KRPMSPPIIP DIPIVDAQSG EDYDSFLMYS DDVLRSPSGS QRPVSDDTE HLVNGRMDFA 1380  
FPGSTNSLHR MTTTAAAYG THLSPHVPHR VLSTSTLTR DYNLSRSEH SHSTTLPRDY 1440  
STLTSVSSH SRITAGVPDT PTRLVFSALG PTLKRVSWQE PRCEPLQGY SVEYQLLNGG 1500  
ELHRLNPNP AQTSVVVEDL LFNHSVVRV RAQSQBGWGR EREGVITIES QVHPQSPLCP 1560  
LPGSAFTLST PSAPGLPVFT ALSPDSLQLS WERPRRPNED IVGYLVTCM AQGGGPATAF 1620  
RVDGSPESR LTVPLGSENV PYKFKVQART TEGFGPEREG ITIESQDGG PFPQLGSRAG 1680  
LFQHPQLQSEY SSITTTHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTQEF VSRTLTTSST 1740  
LSTHMDQQFF QT

Seq ID NO: 562 DNA sequence  
Nucleic Acid Accession #: NM\_013332.1  
Coding sequence: 1..63

85  
1 11 21 31 41 51

GCACGAGGCG GCTTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT 60  
AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC 120  
CGGCTGTTC CCGGAGGGGT CCAGAGGCCT TTCAGAAGGA GAAGGCAGCT CTGTTTCTCT 180  
GCAGAGGAGT AGGGTCTTTT CAGCCATGAA GCATGTGTTG AACCTCTACC TGTAGGTGT 240  
GGTACTGACC CTACTCTCCA TCTTCGTAG AGTGTAGGAG TCCTTAGAAG GCTTACTAGA 300  
GAGCCCATCG CCTGGGAGCTT CCTGGACCAC CAGAAGCCAA CTAGCCAACA CAGAGCCAC 360  
CAAGGCGCTT CCAGACCATC CATCCAGAAG CATGTGATAA GACCTCCTTC CATACTGGCC 420  
ATATTTTGA ACATGACCT AGACATGTCC AGATGGGAGT CCCATTCTTA GCAGACAAGC 480  
TGAGCACCGT TGTAAACAGA GAACTATTAC TAGGCCTTGA AGAACCTGTC TAACTGGATG 540  
CTCATTCGCT GGGCAAGGCC TGTTTAGGCC GGTTCGGGTG GCTCATGCTT GTAATCCTAG 600  
CACTTTGGGA GGCTGAGGTG GGTGGATCAC CTGAGGTGAG GAGTTGAGGA CCAGCCTCGC 660  
CAACATGGCG AAACCCCATC TCTACTAAAA ATACAAAAGT TAGCTGGGTG TGGTGGCAGA 720  
GGCCTGTAAT CCCAGTTTCT TGGGAGGCTG AGGCGGAGGA ATTGCTTGAA CCGGGGACG 780  
GAGGTTGCAG TGAACCGAGA TCGCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC 840  
CATCTCAAAA AAAAAAGAA AAAAAAAGC CTGTTTAATG CACAGGTGTG AGTGGATTGC 900  
TTATGGCTAT GAGTAGGTG GATCTCGCCC TTACCCCGGG GTCTGGTGTA TGCTGTGCTT 960  
TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTCAGCTGT TGGGAGATGG 1020  
TGATATTTTC AACCTACTTT CCTAAACATC TGCTGGGGT TCCTTAGTTC TTGAATGTCT 1080  
TATGCTCAAT TATTTGGTGT TGAGCCTCTC TTCCACAAGA GCTCTCCAT GTTTGGATAG 1140  
CAGTTGAAGA GGTGTGTGG GTGGGCTGTT GGGAGTGAGG ATGGAGTGT CAGTGCCCAT 1200  
TTCTCATTTT ACATTTTAAA GTCGTTCTTC CAACATAGTG TGTATTGGTC TGAAGGGGGT 1260  
GGTGGGATGC CAAAGCCTGC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA 1320  
TTTTTCTAA CTAATAAAGT GGAATATATA TTTCAAAAA AAAAAAAA AA

Seq ID NO: 563 Protein sequence  
Protein Accession #: NP\_037464.1

1 11 21 31 41 51  
MKHVLNLYLL GVVLTLISIP VRVMESLEGL LESPSPGTGW TTRSQLANTE PTKGLPDHPS 60  
RSM

Seq ID NO: 564 DNA sequence  
Nucleic Acid Accession #: NM\_023915.1  
Coding sequence: 250..1326

1 11 21 31 41 51  
GGCAGGAGGG TTTCGTTTTC ATGCTTTACC AGAAAAATCCA CTTCCTTGCC GACCTTAGTT 60  
TCAAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120  
TGGAATGAGC AGCCAGCCAC CACAATGAAA GAAATCAAAC CAGGAATAAC CTATGCTGAA 180  
CCCAAGCCTC AATCGTCCCC AAGTGTCTCC TGACAGCATC CTTCGCTTAC AGTGATCAC 240  
AACTGAAGAA TGGGGTTCAA CTTCAGCTCT GCAAAATTAC CAAATAACGA GCTGCAGCGC 300  
CAAGAGATGC ACAATTCAGG CAACAGGAGC GACGGGCCAG GAAAGAACAC CACCCCTCAC 360  
AATGAATTTG ACACAATTTG CTTCGCGGTG CTTTATCTCA TTATATTGT GGCAAGCATC 420  
TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAAC CAGCTTCATA 480  
TTCTATCTCA AAAACATAGT GGTTCGAGC CTCATAATGA CGCTGACATT TCCATTTCGA 540  
ATAGTCCATG ATGCAGGATT TGGACCTTGG TACTTCAAGT TTATTCTCTG CAGATACACT 600  
TCAGTTTTGT TTTATGCAAA CATGTATACT TCCTTCGTGT TCCTTGGGCT GATAAGCATT 660  
GATGCTATC TGAAGTGTGT CAAGCATTG GGGACTCTC GGATGTACAG CATAACCTTC 720  
ACGAAGGTTT TATCTGTTTG TGTTTGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780  
ATCCTGACAA ATGGTCAGCC AACAGAGGAC AATATCCATG ACTGCTCAAA ACTTAAAGT 840  
CCTTTGGGGG TCAAAATGGC TACGGCAGTC ACCTATGTGA ACAGCTGCTT GTTTGTGGCC 900  
GTGCTGGTGA TTCTGATCGG ATGTTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960  
AGGCAATTC TAACTCACTC AAGCCGAAAG CGAAACATA ACCAGAGCAT CAGGGTGTGT 1020  
GTGGCTGTGT TTTTACCTG CTTTCTACCA TATCACTGT GCAGAATTC TTTTACTTTT 1080  
AGTCACTTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAGAA 1140  
ATTACACTTT TCTGTCTGCT GTGTAATGTT TGCCTGGATC CAATAATTTA CTTTTCATG 1200  
TGTAGGTATC TTTCAGAGG GCTGTTCAAA AAATCAAATA TCAGAACAGG GAGTGAAAGC 1260  
ATCAGATCAC TGAAGTGTG GAGAAGATCG GAAGTTGCGA TATATTATGA TTACTGTAT 1320  
GTGTAGGCTT TTTATTGTTT GTTGAATCG ATATGTACAA AGTGTAAATA AATGTTTCTT 1380  
TTCAATATCC TTAATAAAAA AA

Seq ID NO: 565 Protein sequence  
Protein Accession #: NP\_076404

1 11 21 31 41 51  
MGFNLTLAKL PNNELHGGES HNSGNRSDGP GKNTTLENEF DTIVLPVLYL IIFVASILLN 60  
GLAVWIFPHI RNKTSFIFYL KNIVVADLIM TLTFFPRIVH DAGFGPWYFK PILCRYTSVL 120  
FYANMYTSIV FLGLISIDRY LKVVKPFQDS RMYSTFTKV LSVCVWVIMA VLSLFNIIIT 180  
NQPTEDNIH DCSKLKSLPG VKWHTAVTYV NSCLPVAVLV ILIGCYIAIS RYIHKSSRQF 240  
ISQSSRRKHK NQSIKRVVAV PFTCFPLPYHL CRIPPTFPHL DRLLDESQK ILIYCKEITL 300  
FLSACNVCLD PIIFYFMCRS PSRRLFKKSN ITRSESIRS LQSVRRSEVR IYYDYTDV

Seq ID NO: 566 DNA sequence  
Nucleic Acid Accession #: NM\_005365.1  
Coding sequence: 1..948

1 11 21 31 41 51  
ATGTCCTCTG AGCAGAGGAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA 60  
GAGGACTTGG GCTCATGAGG TGCACAGGAA CCCACAGCGC AGGAGGAGGA GACTACCTCC 120  
TCCTCTGACA GCAAGGAGGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCAGAGT 180  
CCTCAGGAGG GCGCTTCTCT CTCCATTTCC GTCTACTACA CTTTATGGAG CCAATTTCAT 240  
GAGGGCTCCA CGAGTCAAGA AGAGGAAGAG CCAAGCTCCT CGGTGAGCCC AGCTCAGCTG 300  
GAGTTCATGT TCCAAGAAGC ACTGAAATTG AAGGTGGCTG AGTTGGTTCA TTTCTGCTC 360



CACAAATATC GAGTCAAGGA GCCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAAA 420  
 AATTACAAGC GCTACTTTCC TGTGATCTTC GGCAAGCCTT CCGAGTTCAT GCAGGTGATC 480  
 TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCGGGCCACT CCTACATCCT TGCTACTGCT 540  
 CTGGGCTCTT CGTGGGATAG CATGCTGGGT GATGGTCATA GCATGCCCAA GGCCGCCCTC 600  
 CTGATCATTG TCGTGGGTGT GATCCTAACC AAAGACAAC TCGCCCTCGA AGAGGTTATC 660  
 TGGGAAGCGT TGAGTGTGAT GGGGGTGTAT GTTGGGAAG AGCACATGTT CTACGGGGAG 720  
 CCCAGGAAGC TGCTCACCCA AGATTGGGTG CAGGAAAAC ACCTGGAGTA CCGGCAGGTG 780  
 CCCGGCAGTG ATCCTGCACA CTACGAGTTC CTGTGGGGTT CCAAGGCCCA CGCTGAAACC 840  
 AGCTATGAGA AGGTCAATAA TTATTGGGTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC 900  
 CCATCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAGG GAGTCTGA

Seq ID NO: 567 Protein sequence  
 Protein Accession #: NP\_005356.1

1 11 21 31 41 51  
 MSLEQRSPHC KPDEDLEAQQ EDLGLMGAQE PTGEEETTS SSDSKEEEVS AAGSSSPPOS 60  
 PQGGASSIS VYTTLWSQFD EGSSSQEEEE PSSSDVPAQL EFMFQEAALK KVALVHFL 120  
 HKYRVKEPVT KAEMLESVIK NYKRYFPVIF GKASEFMQVI FGTDVKEVDP AGHSYILVTA 180  
 LGLSCDSMLG DGHSMPKAAL LIIVLGVILT KDNCAPEEVI WEALSVMGVY VGKHEMPYGE 240  
 PRKLLTQDWV QENYLEYRQV PGSDPAHYEF LWGSKAHAET SYEKVINYL VMLNAREPICY 300  
 PSLYEVLGE EQEGV

Seq ID NO: 568 DNA sequence  
 Nucleic Acid Accession #: NM\_014400  
 Coding sequence: 86..1126

1 11 21 31 41 51  
 GGTACTCAT CCTGGGCTCA GGTAAGAGGG CCGAGCTCG GAGGCGGCAC ACCCAGGGGG 60  
 GAGCCCAAGG GAGCAGGACG GAGCCATGGA CCCCGCCAGG AAAGCAGGTG CCCAGGCCAT 120  
 GATCTGGACT GCAGGCTGCG TGCTGCTGCT GTGCTTGC GCAGGAGCGC AGGCCCTGGA 180  
 GTGCTACAGC TGCTGTCAGA AAGCAGATGA CGGATGCTCC CGGAACAAGA TGAAGACAGT 240  
 GAAGTGCGCG CCGGGCGTGG ACCTCTGCAC CGAGGCCGTG GGGGCGGTGG AGACCATCCA 300  
 CGGACAATTC TCGCTGCGAC TGCSGGGTG CGGTTCGGGA CTCCCGGCA AGAATGACCG 360  
 CGGCTGGAT CTTCAGGGG TCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420  
 CTGCAACGCC AAGCTCAACC TCACCTCGCG GCGCTCGAC CCGGCAGGTA ATGAGAGTGC 480  
 ATACCGGCCC AACGGCGTGG AGTGCTACAG CTGTGTGGGC CTGAGCCGGG AGGCGTGCCA 540  
 GGGTACATCG CCGCGGTGCG TGAGCTGCTA CAACGCCAGC GATCATGTCT ACAAGGCTG 600  
 CTTCGACGCG AACGTCACTT TGACGCGCAG TAATGTGACT GTGCTCTTGC CTGTCCGGGG 660  
 CTGTGTCCAG GATGAATCT GCACCTCGGA TGGAGTAACA GGGCCAGGTT TCACGCTCAG 720  
 TGCTCTCTGT TGCCAGGGGT CCGCTGTAA CTCTGACCTC CGCAACAAGA CCTACTTCTC 780  
 CCCTCGAATC CCACCCCTTG TCCGGCTGCC CCTCCAGAG CCCACGACTG TGGCCTCAAC 840  
 CACATCTGTC ACCACTTCTA CCTCGGCCCC AGTGAGACCC ACATCCACCA CCAAAACCAT 900  
 GCGAGCGCCA ACCAGTCAGA CTCGAGACA GGGAGTAGAA CAGGAGGCTT CCGGGGATGA 960  
 GGAGCCGAGG TTGACTGGAG GCGCCGCTGG CCACCGAGC CGCAGCAATT CAGGGCAGTA 1020  
 TCCTGCAAAA GGGGGCCCC AGCAGCCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080  
 ATTGGCAGCC CTCTCTGTGG CGTGGCTGCG TGGTGTCTTA CTGTGAGCTT CTCCACCTGG 1140  
 AAATTTCCCT CTCACTACTT TCTCTGGCCC TGGGTACCCC TCTTCTCATC ACTTCTCTGT 1200  
 CCCACCACTG GACTGGGCTG GCCCAGCCCC TGTTTTCCA ACATTCCTCA GTATCCCTCAG 1260  
 CTCTCTGCTG GCTGTTTGG GCTTTTGGGA AATAAATAAC CGTGTATAT ATCTGCTCAG 1320  
 GGGTGTCTTA GCTTTTGGG GACAGCTCTT GTATCTCTCT CATCCTTGTG TCTCCGCTTG 1380  
 TCCTCTGTG ATGTAGGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGGA GGTGAGAGAG 1440  
 AGGATGCTAA GCTTCTACT CACTTCTCC TAGCCAGCTT GGAATTGGA GCGTGGGGTG 1500  
 GGTGGGACAA TGGTCCCCA CTCTAAGCAC TGCCCTCCCC ACTCCCGCA TCTTTGGGGA 1560  
 ATCGGTCCCC CATATGCTT CTTACTAGA CTGTGAGCTC CTCGAGGGA GGGACCGTGC 1620  
 CTTATGCTG TGTGTGATCA GTTCTGGCA CATAAATGCC TCAATAAGA TTTAATTACT 1680  
 TTGTATAGTG AAAAAAA

Seq ID NO: 569 Protein sequence  
 Protein Accession #: NP\_055215

1 11 21 31 41 51  
 MDPARKAGAQ AMIWTAGWLL LLLLRGGAQA LECYSCVQKA DDGCSNPKMK TVKCAPGV DV 60  
 CTEAVGAVET IHGQFSLAVX GCGSGLPGKN DRGLDLHGLL AFIQLQCAQ DRCNAKLNLT 120  
 SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPFVVS CYNASDHVYK GCFDGNVTLT 180  
 AANVTVSLPV RGCVDDEFCT RDGVTGPFT LSGSCQGSR CNSDLRNKTY FSPRIPLVLR 240  
 LPPPEPTTVA STTSVTSTST APVRPTSTTK PMPAPTSQTP RQGVHEEASR DEEPRLTGGA 300  
 AGHQDRNSNG QYPAKGFPQQ PHNKGCVAPT AGLAALLLAV AAGVLL

Seq ID NO: 570 DNA sequence  
 Nucleic Acid Accession #: NM\_005329.1  
 Coding sequence: 1..1662

1 11 21 31 41 51  
 ATGCCGGTGC AGCTGACGAC AGCCCTGGGT GTGGTGGGCA CCAGCCTGTT TGCCCTGGCA 60  
 GTGCTGGGTG GCATCCTGGC AGCCTATGTG ACGGGCTACC AGTTCATCCA CACGAAAAG 120  
 CACTACCTGT CCTTGGGCTT GTACGGGCGC ATCCTGGGCC TGACCTGCTT CATTGAGAGC 180  
 CTTTTCGCTT TCTTGGAGCA CCGCGCATG CGAGTGCTCG GCCAGGCCCT GAAGCTGCCC 240  
 TCCCGCGCGC GGGGCTCGGT GGCAGTGTGC ATTGCCGCGT ACCAGGAGGA CCTGACTAC 300  
 TTGCGCAAGT GCCTGCGCTC GCGCCAGCGC ATCTCCTTCC CTGACCTCAA GGTGGTCACT 360  
 GTGTGTGATG GCAACGCCCA GGAGGACGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420  
 GCGGCGACCG AGCAGGCCGG CTTCTTGTG TGGCGCAGCA ACTTCCATGA GGCAGGGCAG 480  
 GGTGAGACGG AGGCCAGCCT GCAGGAGGCG ATGGACCGTG TGCGGATGT GGTGCGGGCC 540  
 AGCACCTTCT CGTGATCAT CATGAAGTGG GAGGCAAGC GCGAGGTCAT GTACACGGCC 600



TTCAAGGCC TCGGCGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660  
 GATCCAGCCT GCACCATCGA GATGCTTCCA GTCCCTGGAG AGGATCCCCA AGTAGGGGGA 720  
 GTGCGGGGAG ATGTCCAGAT CCTCAACAG TACGACTCAT GGATTTCCCTT CCTGAGCAGC 780  
 GTGCGGTACT GGATGGCCTT CAACGTGGAG CGGGCTGCCC AGTCTACTT TGGCTGTGTG 840  
 CAGTGTATTA GTGGGCCCTT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCTCGGAG 900  
 GACTGGTACC ATCAGAAATT CCTAGGCAGC AAGTGCAGCT TCGGGGATGA CCGGCACCTC 960  
 ACCAACCGAG TCCTGAGCCT TGGCTACCGA ACTAAGTATA CCGCGGCTC CAAGTGCTCTC 1020  
 ACAGAGACCC CCACTAAGTA CCTCCGGTGG CTCAACCAGC AAACCCGCTG GAGCAAGTCT 1080  
 TACTTCCGGG AGTGGCTCTA CAACTCTCTG TGGTTCCATA AGCACCACCT CTGGATGACC 1140  
 TACGAGTCAG TGGTCAACGG TTCTTCCCTT TTCTTCCCTA TTGCCAOGGT TATACAGCTT 1200  
 TTCTACCGGG GCGCATCTG GAACATTCTC CTCTTCCCTG TGAACGGTGA GCTGGTGGGC 1260  
 ATTATCAAGG CCACTACGCG CTGCTTCCCTT CCGGGCAATG CAGAGATGAT CTTCATGTCC 1320  
 CTCTACTCCC TCCTCTATAT GTCCAGCCTT CTGCGGSCCA AGATCTTTGC CATTGTCTACC 1380  
 ATCAACAAAT TGGGCTGGGG CACCTCTGGC CGAAAAACCA TTGTGGTGAA CTCATTGGC 1440  
 CTCATTCTCG TGTCCATCTG GGTGGCAGTT CTCTGCGGAG GGCTGGCCTA CACAGCTTAT 1500  
 TGCCAGGACC TGTTCAGTGA GACAGAGCTA GCCTTCCCTG TCTCTGGGGC TATACTGTAT 1560  
 GGCTGCTACT GGTGGCCCTT CCTCATGCTA TATCTGGCCA TCATCGCCCG GCGATGTGGG 1620  
 AAGAAGCCGG AGCAGTACAG CTTGGCTTTT GCTGAGGTGT GA

Seq ID NO: 571 Protein sequence  
 Protein Accession #: NP\_005320.1

1 11 21 31 41 51  
 MPVQLTTALR VVGTSLFALA VLGGILAAVY TGYQFIHTEK HYLSPGLYGA ILGLHLLIQS 60  
 LPAPLEHRRM RRAGQALKLP SPRRGVALC IAAYQEDFDY LRKCLRSAQR ISFPDLRVVM 120  
 VVDGNRQEDA YMLDIFHEVL GTEGAGFFV WRSNFHEAGE GETEASLQEG MDRVRDVVRA 180  
 STFCIMQKRW GGRKRVMTYA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEEDQVQGG 240  
 VGGDVQILNK YDSWISFLSS VRYWMAFNVE RACQSYFGCV QCISGFLPMY RNSLLQQFLE 300  
 DWYHKPLGKS KCSFGDDRHL TNRVLSLGYR TKYTARSKCL TETPTKYLW LNQQTRWSKS 360  
 YPREWLYNSL WPHKHLHML YESVVTGFFP FPLIATVIQL FYRGRINWL LFLTLVQLVG 420  
 IIKATYACFL RGNAMIFMS LYSLLYMSL LPAKIPAIAT INKSGWGTSG RKTIVNVFIG 480  
 LIPVSIWVAV LLGGLAYTAY QDLFSETEL AFLVSGAILY GCYNVALML YLAIIARROG 540  
 KKPEQYSLAP AEV

Seq ID NO: 572 DNA sequence  
 Nucleic Acid Accession #: E05 sequence  
 Coding sequence: 148-7095

1 11 21 31 41 51  
 CACACATACG CACGCAAGAT CTCACTTCCA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAAC ACTTCTCTCG CTCCCTCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGCGGAGGGG CGCAGAGCCG TCTGGAAATG CGAATCTCAA AGCGTTTCCT CGCTTGCAAT 180  
 CAGCTCCTCT GTGTTTGCGG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTGTGTTAAG AGATTGGCTG GTCCCTATACA GGAGCACTGA ATCAAAAAAA TTGGGGAAAG 300  
 AATATCCAA CTGATTAATAG CCAAAAACAA TCTCTATCA ATATTGATGA AGATCTTACA 360  
 CAGTAAATG TGAATCTTAA GAAACTTAAA TTTAGGGTT GGGATAAAC ATCATTTGAA 420  
 AACACATTCA TTCATAACAC TGGGAAAAACA GTGAAATTA ATCTCACTAA TGACTACCGT 480  
 GTCAGCGGAG GAGTTTCAGA AATGGTGTIT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AATGCAATA TGTATCTGA TGGATCAGAG CATAGTTTAG AAGGACAAAA ATTTCCACTT 600  
 GAGATGCAAA TCTACTGCTT TGATGOGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAAGGGGA AGTTAAAGAG TTTATCCATT TTGTTTGAGG TTGGGACAGA AGAAAAATTTG 720  
 GATTTCAAAG CGATTATATGA TGGAGTCGAA AGTGTAGTCT GTTTTGAGGA GCAGGCTGCT 780  
 TTAGATCCAT TCATACTGTT GAACCTCTCG CCAACTCAA CTGACAAGTA TTACATTAC 840  
 AATGGCTCAT TGACATCTCC TCCCTGCACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTATGCA TCTCTGAAG CCAGTTGGCT GTTTTTTGTT AAGTTCTTAC AATGCAACAA 960  
 TCTGTTATG TCATGCTGAT GGACTACTTA CAAAAAATT TTGAGAGCA ACAGTACAAG 1020  
 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTCTAGA AGCAGTTTGT 1080  
 AGTTCCAGAC CAGAAAAATG TCAGGCTGAC CCAGAGAATT ATACCAGCTT TCTGTTTACA 1140  
 TGGGAAAGAC CTGAGTCTGT TTATGATACC ATGATTGAGA AGTTTGCACT TTTGTACCAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATTC TCAATAATTT GCTACCCAAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCACTAATG GCTTATATGG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380  
 AATCCTGAAC TTGATCTTTT CCTGAAATTA ATTGGAACCT AAGAAATAAT CAAGGAGGAG 1440  
 GAAGAGGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGGTAGAGA CAGTGTCTACA 1500  
 AACCAAAATCA GGAAGAAAGGA ACCCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACAGAAATCA ATGAAGCCAA GACTAACCGA TCCCCAACAA GAGGAAGTGA ATTTCTGGA 1620  
 AAGGGTGATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680  
 ACAGAAAAAG ATATTTCCTT GACTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGTACTTT CAGCCTCTTT AAATGATGGC TCTAAAACCTG TTCTTAGATC TCCACATATG 1800  
 AACTTGTGGG GGACTGCGAA ATCCTTAAAT ACAGTTTCTA TAACAGAAAT TGAGGAGGAG 1860  
 AGTTTATGGA CCAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCCAGTCCC 1920  
 GCAACTTCTG CTATCCCATT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTTCTCTC 1980  
 GAAAAACCCAG AGACAATAAC ATATGATGTC CTTATACCAG AATCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTATCAGG TTCAGAAAGG TCACATAAGG ATCCTTCTAT GGAGGGAAT 2100  
 GTGTGGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTTCTCC AGCTTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAAACCAAG 2220  
 TCCTTTCTG CAGGCCCACT GATGTACAG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280  
 CATTATCTA CTTTGGCTA CTTCCCACT GAGGTAAACAC CTCATGCTTT TACCCATCC 2340  
 TCCAGACAAC AGGATTGGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAACCG 2400  
 GTATACAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTCTAGTC 2460  
 ACCCTTTTGT TGCTTGACAA TCAGATCCTC AACACTACCC CTGCTGCTTC AAGTAGTGAT 2520  
 TCGGCTTGGC ATGCTACGCC TGTATTTCCC AGTGTGATG TGTCAATTGA ATCCATCCTG 2580  
 TCTTCTATG ATGGTGCAAC TTTGCTTCCA TTTTCTCTG TCTCTTCAG TAGTGAATTG 2640  
 TTTGCCATC TGCATACAGT TTCTCAAATC CTTCCACAAG TTACTTCAGC TACCGAGAGT 2700  
 GATAAGGTGC CCTTGATGTC TTCTCTGCCA GTGGCTGGGG GTGATTGCT ATTAGAGCCC 2760

	AGCCTTCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTGGTAGTG	AATCTGGTGT	TCITTTATAA	ACGCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTGATG	CCATGATGCA	TGCACGTTCT	TCAGGGCCTG	AACTTCTTTA	TGCTTGTCT	2940
5	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTTTCTTACA	GTTCTGCAAT	ACCTGTGCAT	3000
	GATTCTGTGG	GTGTAACCTA	TCAGGGTTCC	TTATTTAGCG	GCCTAGCCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCCCTCT	3120
	GGTGATGGGG	AATGGTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACTT	GTTTCTGTAG	CTGAATTATC	ATATACAACA	3240
10	TCGTGTGTTG	GTGATGATAA	TAAGGCGCTT	TCTAAAAGTG	AAATAATATA	TGGAATGAG	3300
	ACTGAAGTGC	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCAAT	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTGT	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGCGAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACTTTTTCG	TTCAAACTAC	ACATACCTGC	3540
15	TCCTCAAGCAT	CTGGTGACAC	TTCCCTTAAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATTT	CTACAACCTT	CCITTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAAACCTG	TCTTCCAGCT	GTGCCCAGTG	ATCCAATATT	GGTGTAAACC	3780
	CCCAAGTTTG	ATAAAATTAG	TTCTACAATG	TTGCATCTCA	TTGTATCAAA	TTCTGCTTCA	3840
20	AGTGAACA	CTGCTCACTG	TACATCTGTA	CCAGTTTTTG	ATGTGTGCGC	TACTTCTCAT	3900
	ATGCACTCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTCTAT	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTATA	AAAGTGAAG	TTCCCAACCA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAAC	CAGGCCCATC	CCCCAAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATGTAGTAA	CCATTAAATA	CACATAATAA	TAAGCTTATA	4140
25	CATTCGATG	AAATTTAATC	CTCCACCAAA	AGTTCTGTTA	CTGTAAGGT	ATTGTCTGTT	4200
	ATTCCACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTCGTGTTT	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCTGTAAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
30	AGAGGTAGTG	ATGGCTTATC	CATTCTAAG	TGTATGTCAT	GCTCATCTTA	TAGAGAAATCA	4500
	CAGGAAAAGG	TAATGAATGA	TTCAAGACAC	CAAGAAAACA	GTCTTATGGA	TCAGAAATAT	4560
	CCAATCTCAT	ACTCACTATC	TGAGAAATCT	GAAGAAAGTA	ATAGAGTCAC	AAGTGTATCC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAAGT	CCTGTAAAT	CAACATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACTGGTAG	TGCTCTGCTT	4740
35	CCTCTCAGCC	CTGAATCTAA	AGCATGGGCA	GTTCTGACAA	GTGATGAAGA	AAGTGGATCA	4800
	GGGCAGAGTA	CCTCAGATAG	CCTTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTGTGA	4860
	GACACTAATG	AAAAGTGGT	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	AAATACTCCT	4920
	GGATTCCAC	AGTCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAAAT	GTTCCACGTT	4980
	TCAGAGGCG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	5040
40	GAATCCGAGA	AGAAGGCGAT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	5100
	CTAGTGGTTC	TTGTTGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAAACCC	TATCTTTCCA	5220
	ATTTCAAGATG	ATGTCGAGC	AAITCCAATA	AAGCACTTTC	CAAGCATGTT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGACACTGA	AAGAGTTTTA	CCAGGAAGTG	5340
45	CAGAGCTGTA	CTGTGACTTT	AGGTATTACA	GCAGACAGCT	CCAACCAACC	AGACAACAG	5400
	CACAAGATTC	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460
	CTTGTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTTATG	TGATGGCTAC	5520
	AACAGACCAC	AAGCTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTT	5580
50	TGGAGAATGA	TATGGGAACA	TAATGTGGAA	GTTATTGTCA	TGATAACAAA	CCTCGTGGAG	5640
	AAAGGAAGGA	GAAAATGTGA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTACGGGAAC	5700
	TTTCTGGTCA	CTCAGAGAGG	TGTGCAAGTG	CTTGCTTATT	ATAGTGTGAG	GAATTTTACT	5760
	CTAAGAAACA	CAAAAATAAA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	5820
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	5880
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGGCCCATG	CAGTGGGGCC	TGTTGTGCTC	5940
55	CACGTGAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTG	TGCTAGACAG	TATGTTGCAG	6000
	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAACACAT	CCGTTTCAAA	6060
	AGAAATTTAT	TGTTCAAAAC	TGAGGAGCAA	TATGCTTCTA	TTCAATGATC	ACTGGTTGAG	6120
	GCCATACTTA	GTAAAGAAAC	TGAGGTGCTG	GACAGTCTAT	TTCAATGCTA	TGTTAATGCA	6180
	CTCCTCATTC	CTGGACAGGC	AGGCAAAACA	AAGCTAGAGA	AACAATTTCA	GCTCCTGAGC	6240
60	CAGTCAATA	TACAGCAGAG	TGACTATTCT	GCAGCCCTAA	AGCAATGCAA	CAGGGAAGAG	6300
	AATCGAAGTT	CTTCTATCAT	CCCTGTGGAA	AGATCAAGGG	TTGGCATTTT	ATCCCTGAGT	6360
	GGAGAAGGCA	CAGACTACAT	CAATGCCTCC	TATATCATGG	GCTATTACCA	GAGCAATGAA	6420
	TTCATCATTA	CCAGCAGCCC	TCTCCTTCAT	ACCATCAAGG	ATTTCTGGAG	GATGATATGG	6480
	GACCATAATG	CCCACTGGT	GGTTATGATT	CCTGATGGCC	AAAACATGGC	AGAAGATGAA	6540
65	TTTGTTTACT	GGCCAAATAA	AGATGAGCCT	ATAAATTTGT	AGAGCTTTAA	GGTCACTCTT	6600
	ATGGCTGAAG	AACCAAAATG	TCTATCTAAT	GAGGAAAAC	TTATAATTCA	GGACTTTATC	6660
	TTAGAAGCTA	CACAGGATGA	TTATGTACTT	GAAGTGAGGC	ACTTTCAGTG	TCCTAAATGG	6720
	CCAAATCCAG	ATAGCCCCAT	TAGTAAAACT	TTTGAACCTA	TAGTGTGTTT	AAAAGAAGAA	6780
	GCTGCCAATA	GGGATGGGCC	TATGATTGTT	CATGATGAGC	ATGGAGGAGT	GACGGCAGGA	6840
70	ACTTCTGTG	CTCTGACAC	CCTTATGCAC	CAACTAGAAA	AAGAAAATTC	CGTGATGTTT	6900
	TACCAGGTAG	CCAAGATGAT	CAATCTGATG	AGGCCAGGAG	TCTTGTCTGA	CATTGAGCAG	6960
	TATCAGTTTC	TCTACAAAGT	GATCCTCAGC	CTTGTGAGCA	CAAGGCAGGA	AGAGAATCCA	7020
	TCCACCTCTC	TGGACAGTAA	TGGTGACGCA	TTGCCCTGAT	GAATATAGC	TGAGAGCTTA	7080
	GAGTCTTTAG	TTTAAACACG	AAAGGGGTGG	GGGGAAGTAC	ATCTGAGCAT	TGTTTCTCTC	7140
75	TTCTTAAAT	TAGGACAGAA	AATCAGTCTA	GTTCTGTTAT	CTGTTGATTT	CCCATCAGCT	7200
	GACAGTAACT	TTATGACAT	AGGATTCTGC	CGCCAAATTT	ATATCATTA	CAATGTGTGC	7260
	CTTTTTCGAA	GACTTGTAAAT	TTACTTATTA	TGTTTGAAT	AAAATGATTG	AAATTTTACG	7320
	TATTTCTAAG	AATGGAATTT	TGGTATTTTT	TTCTGTATTG	ATTTTAAACG	AAAATTTCAA	7380
	TTTATAGAGG	TTAGGAATTC	CAAACTACAG	AAAAATGTTT	TTTTTAGTGT	CAAAATTTTA	7440
80	GCTGTATTTG	TAGCAATTTAT	CAGGTTTGCT	AGAAATATAA	CTTTTAAATC	AGTAGCCTGT	7500
	AAATAAAACA	CTCTTCCATA	TGATAATCAA	CATTTTACAA	CTGCAGTATT	CACCTAAAGT	7560
	AGAAATAATC	TGTTACTTTAT	TGTAATAACT	GCCCTAGTGT	CTCCATGGAC	CAAAATTTATA	7620
	TTTATAATTG	TAGATTTTTA	TATTTTACTA	CTGAGTCAAG	TTTCTAGTGT	CTGTGTAATT	7680
	GTTTAGTTTA	ATGACGTAGT	TCATTAGCTG	GTCTTACTCT	ACCAGTTTTT	TGACATTTGA	7740
85	TTGTGTTACC	TAAGTCATTA	ACTTTGTTTC	AGCATGTAAAT	TTTAACTTTT	GTGGAAAATA	7800
	GAATACACCT	CATTTTGAAA	GAAAGTTTTT	TGAGAAATAA	ACCTTACCAA	ACATTGTTCA	7860
	AATGGTTTTT	ATCCAGGGAA	TTGCAAAAT	AAATATAAAT	ATTGCCATTA	AAAAAAAATA	7920
	AAAAAAAATA	AAAAAAAATA	AAAA				

Seq ID NO: 573 Protein sequence:  
Protein Accession #: Eos sequence

5	1	11	21	31	41	51	
	MRILKRFLAC	IQLLCVCRLD	WANGYYRQQR	KLVEEIGWSY	TGALNQKNWG	KKYPTCNSPK	60
	QSPINIDEDL	TQVNVNLKKL	KPQGWDKTSL	ENTFIHNTGK	TVEINLTNDY	RVSQGVSEMV	120
10	FKASKITPHW	GRCNMSSDGS	EHSLEGQKFP	LEMQIYCFDA	DRPSSFREAV	KGKGKLRALS	180
	ILFEVGTSEN	LDFKAIIDGV	ESVSRFGKQA	ALDPPILLNL	LPNSTDKYYI	YNGSLTSPPC	240
	TDTVDWIVFK	DTVSISSBSQ	AVFCVLTMQ	QSGYVLMMDY	LQNNFRQQY	KFSRQVPSY	300
	TGKEEIEHAV	CSSEPENVOA	DPENYTSLLV	TWERPRVVDY	TMIEKFAVLY	QQLDGEDQTK	360
	HEFLTDGYQD	LGAILNNLLP	NMSYVLQIVA	ICTNGLYQKY	SDQLIVDMPT	DNPELDLFE	420
15	LIGTEBIIKE	EEEGKDIEEG	AIVNPGRDSA	TNQIRKKEPQ	ISTTHYNNRI	GTYNEAKTN	480
	RSPTRGSEFS	KGKDVNPTSL	NSTSQPVTKL	ATEKDISLTS	QTVTELPPT	VEGTSASLND	540
	GSKTULRSFH	MNLSSTAESL	NTVSIITEYEE	ESLLTSFKLD	TGAEDSSGSS	PATSAIPFIS	600
	ENISGGVIPS	SENPEITTYD	VLIPESARNA	SEDSTSSGSE	ESLKDPSMEG	NWFFPSSTDI	660
	TAQPDVSGSR	ESFLQNTYH	IRVDESEKTT	KSFSAGPVMS	QGPSVTDLEM	PHYSTFAYFP	720
20	TEVTPHAFPT	SSRQDQLVST	VNVVYSQTTQ	PVYNGETPLQ	PSYSSEVFPL	VTPLLLDNQI	780
	LNTTPAASSS	DSALHATPVF	PSVDVSFESI	LSSYDGAPLL	PFSSASFSSS	LFRHLHTVSQ	840
	ILFQVTSATE	SKDKVHLASL	PVAGEDLLLE	PSLAQYSDVL	STTHAASETL	EFGSESGVLY	900
	KTLMFSGVEP	PSDSAMMHAR	SSGPEPSYAL	SDNEGSOHIP	TVSYSSAIPV	HDSVGVTYQG	960
	SLFSPGSHIP	IPKSLITPT	ASLLQPTHAL	SGDGEWSGAS	SDSEFLLPDT	DGLTALNISS	1020
25	FVSVAEFTYT	TSVFGDDNKA	LSKSEIYGN	ETELQIPSPN	EMVYPSBSTV	MFNMYDNVKN	1080
	LNASLQETSV	SISSTKGMPF	GSIAHTTTKV	FDHEISQVPE	NNPSVQPTH	VSQASGDTSL	1140
	KPVLSANSEP	ASSDPASSE	LSPSTQLLPY	ETSASFSTEV	LQPSQFQASD	VDTLTKTVLP	1200
	AVPSDPLIVE	TKVDKISST	MLHLIVNSA	SSENMHLSTS	VPVFDVSPTS	EMHGSASLQL	1260
	TISYASEKYE	PVLLKSESSG	QVPSLYSND	ELFQTANLEI	NQAEHPKGRH	VFATPVLSID	1320
30	EPLNLYLINK	IHSDEILTST	KSSVTGKVFA	GIPTVASDTP	VSTDHVSFIG	NGHVAITAVS	1380
	PHRDGTSVST	KLFPSPKATS	ELSHSAKSDA	GLVGGGSDGD	TDDGDDDDDD	DRGSDGLSIH	1440
	KCMSCSYRE	SQEKVMNDS	THENSAMDQN	NPISYSLSEN	SEEDNRVTSS	SDDSGTQMDR	1500
	SPGKSPSANG	LSQKNDGKE	ENDIQTSAL	LPLSPESKAW	AVLTSDEESG	SGQQTSDSLN	1560
	ENETSSTDFSF	ADTKNDGSEI	ILAAAGDSEIT	PGFPOSPSTSS	VTSENSEVFH	VSEAEASNSS	1620
35	HESRIGLAEG	LBSEKKAIVP	LVIVSALTPI	CLVVLVGILI	YWRKCPQTAH	FYLEDSTSPR	1680
	VISTPPTPIF	PISSDDVGAIP	IKHFPKHVAD	LHASSGFTTE	FETLKEFYQE	VQSCVTDLGI	1740
	TADSSNHPND	KHRTSYINIV	AYDHSRVKLA	QLAEKDGKLT	DYINANYVDG	YNRKAYIAA	1800
	QGPLKSTAEF	FWRMIEHNV	EVIVMITNLV	EKGRRKCDQY	WPADGSEBYG	NFLVTQKSVQ	1860
	VLAYTYVRNF	LRNTKIKKG	SQKGRPSGRV	VTQYHYTQWP	DMGVPEYSLP	VLTFVRKAAY	1920
40	AKREAVGPVV	VECSAGVORT	GTIIVLDSML	QQIQHEGTVN	IFGFLKHIRS	QRNVLVQTEE	1980
	QYVFIHDTLV	EAILSKETEV	LDSHIHAYVN	ALLIPGPAGK	TKLEKQPQLL	SQSNIQSDY	2040
	SAALKQCNRE	KHRTSSIIIP	ERSRVGISSL	SGEGTDYINA	SYIMGYQSN	EPITQHPILL	2100
	HTIKDFWRMI	WDHNAQLVVM	IPDQGNMABD	EPVYWNKDE	PINCSEPKVT	LMABEHKCLS	2160
	NEEKLIQDPF	ILEATODDYF	LEVRHPQCPK	WPNPDSPISK	TFELISVIKE	EAANRDGPMI	2220
45	VHDEHGGVTA	GTFCALITLM	HQLEKENSVD	VYQVAKMINL	MRPGVPADIE	QYQFLYKVL	2280
	SLVSTRQSEN	PSTSLDSNGA	ALPDGNIAES	LESILV			

Seq ID NO: 574 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 148-4518

50	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCACTTCGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCOCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
55	CGGCGAGGGG	CGGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAIT	180
	CAGCTCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTGAAG	AGATTGGCTG	GTCCATATACA	GGAGCACTGA	ATCAAAAAAA	TGGGGGAAAG	300
	AAATATCCAA	AGCTTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAATG	TGAATCTTAA	GAACTTAA	TTTCAGGGTT	GGGATAAAAC	ATCATGGGAA	420
60	AACACATTCA	TTTATAACAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTACGCGGAG	GAGTTTCAGA	AATGGTGT	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTATCTCGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCACATT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTGGAGGA	AGCAGTCAAA	660
	GGAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAAG	TTGGGACAGA	AGAAAAATTG	720
65	GATTTCAAAG	CGATTATTGA	TGGAGTGGAA	AGTGTTAGTC	GTTTGGGGA	GCAGGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAAG	CCAGTTGGCT	GTTTTTTGTG	AAGTTCCTAC	AATGCAACAA	960
	TCTGGTTATG	TCATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
70	TCTCTAGAC	AGGTGTTTTT	CTCATACACT	GGAAAGGAAG	AGATTCTATG	AGCAGTTTGT	1080
	AGTTCAGAAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
75	TGCATTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAACTG	AAGAAATAAT	CAAGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGTTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCAACA	CACACTACAA	TGCGATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
80	AAGGTTGATG	TCTCCAAATC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACTGCCACC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAACTG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTGCG	GGACTGCGAG	ATCCTTAAAT	ACAGTTTCTA	TAACAGATAA	TGAGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
85	GCAACTCTTG	TGCTCCCAAT	CATCTCTGAG	AACATATCCC	AAGGATATAT	ATTTTCTCTC	1980
	GAAAAACCCAG	AGACAATAAC	ATATGATGTC	CTTATACCAG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCATCAGG	TTTCAAGAAA	TCATAAAGG	ATCCTTCTAT	GGAGGAAAT	2100

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

GTGTGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTTCTCC AGACTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACAACCAAG 2220  
 TCCTTTTCTC CAGGCCCACT GATGTACACG GGTCCCTCAG TTACAGATCT GGAATGCCA 2280  
 CATTATCTA CCTTTGGCTA CTCCCAACT GAGGTAACAC CTCATGCTTT TACCCCATCC 2340  
 TCCAGACAAC AGGATTTCGT CTCCACGGTC AACGTGGTAT ACTCGCAGAC AACCCAAACG 2400  
 GTATACAAATG CAGAGGCCAG TAATAGTAGC CATGAGTCTC GTATTGGTCT AGCTGAGGGG 2460  
 TTGGAATCCG AGAAGAAGGC AGTTATACCC CTGTGATCG TGTCAGCCCT GACTTTTATC 2520  
 TGTCTAGTGG TTCTTGTGGG TATTCTCATC TACTGGAGGA AATGCTTCCA GACTGCACAC 2580  
 TTTTACTTAG AGGACAGTAC ATCCCTTAGA GTTATATCCA CACCTCCAAC ACCTATCTTT 2640  
 CCAATTTTCAG ATGATGTCCG AGCAATTTCA ATAAAGCACT TTCCAAAGCA TGTTCAGAT 2700  
 TTACATGCAA GTAGTGGGTT TACTGAAGAA TTTGAGACAC TGAAAGAGTT TTACAGGAA 2760  
 GTGCAGAGCT GTACTGTTGA CTTAGGTATT ACAGCAGACA GCTCCAACCA CCCAGACAAC 2820  
 AAGCACAAGA ATCGATACAT AAATATCGTT GCCTATGATC ATAGCAGGGT TAAGCTAGCA 2880  
 CAGCTTGCTG AAAAGTAGG CAAACTGACT GATTATATCA ATGCCAATTA TGTGTATGCC 2940  
 TACAACAGAC CAAAAGCTTA TATTGCTGCC CAAGGCCAC TGAAATCCAC AGCTGAAGAT 3000  
 TTCTGGAGAA TGATATGGGA ACATAATGTG GAAGTTATTG TCATGATAAC AAACCTCGTG 3060  
 GAGAAAGGAA GGAGAAATAT TGATCAGTAC TGCCCTGCCG ATGGGAGTGA GGAGTACGGG 3120  
 AACTTTCTGG TCACCTAGAA GAGTGTGCAA GTGCTTGCTT ATTACTGTG GAGGAATTTT 3180  
 ACTCTAAGAA AGCAAAAAT AAAAAAGGGC TCCAGAAAG GAAGACCCAG TGGACGTGTG 3240  
 GTCAACAGT ATCACTACAC GCAGTGGCCT GACATGGGAG TACCAGAGTA CTCCCTGCCA 3300  
 GTGCTGACCT TTGTGAGAAA GGCAGCCTAT GCCAAGCGCC ATGCAGTGGG GCCTGTGTGC 3360  
 GTCCACTGCA GTGCTGGAGT TGAAGAACA GGCACATATA TTGTGCTAGA CAGTATGTTG 3420  
 CAGCAGATTC AACAGAAAG AACTGTCAAC ATATTTGGCT TCTTAAACA CATCCGTTCA 3480  
 CAAAGAAAT ATTGTGTACA AACTGAGGAG CAATATGTCT TCATTCTATG TACACTGGTT 3540  
 GAGGCGCATC TTAGTAAGA AACTGAGGTG CTGGACAGTC ATATTCATGC CTATGTTAAT 3600  
 GCATCTCTCA TTCTGGACC AGCAGGCAAA ACAAGCTAG AGAAACAATT CCAGCTCCTG 3660  
 AGCCAGTCAA ATATACAGCA GAGTGACTAT TCTGCAGCCC TAAAGCAATG CAACAGGGAA 3720  
 AAGAATCGAA CTCTCTCTAT CATCCCTGTG GAAAGATCAA GGGTTGGCAT TTCATCCCTG 3780  
 AGTGAGAGAG GCACAGACTA CATCAATGCC TCCTATATCA TGGGCTATTA CCAGAGCAAT 3840  
 GAATTCATCA TTACCCAGCA CCTCTCCTT CATACCATCA AGGATTTCTG GAGGATGATA 3900  
 TGGGACATA ATGCCCACT GGTGGTTATG ATTCTGATG GCCAAAACAT GGCAGAAGAT 3960  
 GAATTTGTTT ACTGGCCAAA TAAAGATGAG CCTATAAAT GTGAGAGCTT TAAGGTCACT 4020  
 CTATAGGCTG AAGAACAACA ATGTCTATCT AATGAGGAAA AACTTATAAT TCAGGACTTT 4080  
 ATCTTAGAAG CTACACAGGA TGATTATGTA CTGAAAGTGA GGCACCTTCA GTGCTCTAAA 4140  
 TGGCCAAATG CAGATAGCCC CATTAGTAAA ACTTTTGAAC TTATAAGTGT TATRAAAGAA 4200  
 GAAGCTGCCA ATAGGGATG GCTATGATT GTTCATGATG AGCATGGAGG AGTGACGGCA 4260  
 GGAACCTTCT GTGCTCTGAC AACCTTATG CACCAACTAG AAAAGAGAAA TTCCGTGGAT 4320  
 GTTACACAGG TAGCCAGAT GATCAATCTG ATGAGGCCAG GAGTCTTTGC TGACATTGAG 4380  
 CAGTATCAGT TTCTCTACAA AGTGATCCTC AGCCTTGTTG GCACAAGGCA GGAAGAGAT 4440  
 CCATCCACCT CTCTGGACAG TAATGGTGCA GCATTGCGCT ATGGAAATAT AGCTGAGAGC 4500  
 TTAGAGTCTT ATGTTTAAAC CAGAAAGGGG TGGGGGAGT CACATCTGAG CATGTTTTC 4560  
 CTCTTCTTAA AATTAGGCAG GAAATCAGT CTAGTTCTGT TATCTGTGTA TTTCCCATCA 4620  
 CCTGACAGTA ACTTTCATGA CATAGGATTC TGCCGCCAAA TTTATATCAT TAACAATGTG 4680  
 TGCTTTTTC CAAGACTTGT AATTACTTAA TTATGTTTGA ACTAAAATGA TTGAATTTTA 4740  
 CAGTATTTCT AAGAATGGAA TTGTGTTATT TTTTCTGTA TTGATTTTAA CAGAAAATTT 4800  
 CAATTTATAG AGGTTAGGAA TTCCAACTA CAGAAAATGT TTGTTTTAG TGTCAAATTT 4860  
 TTAGCTGTAT TTGTAGCAAT TATCAGGTTT GCTAGAAATA TAACTTTTAA TACAGTAGCC 4920  
 TGTAAATAAA ACACCTCTCC ATATGATATT CAACATTTTA CAACCTGAGT ATTCACCTAA 4980  
 AGTAGAAATA ATCTGTTACT TATTGTAAAT ACTGCCCTAG TGTCTCCATG GACCAAAATT 5040  
 ATATTATATA TTGTAGATT TTATATTTTA CTACTGAGTC AAGTTTCTTA GTTCTGTGTA 5100  
 ATGTTTATG TTAATGACGT AGTTCATTAG CTGGTCTTAC TCTACCAAGT TTCTGACATT 5160  
 GTATTGTGTT ACCTAAGTCA TTAACCTTGT TTCAGCATGT AATTTTAACT TTTGTGGAAA 5220  
 ATAGAAATAC CTCATTTTGA AAAGAAGTTT TTATGAGAAAT AACACCTTAC CAAACATTGT 5280  
 TCAATGGTT TTTATCCAAG GAATTGCAAA AATAAATATA AATATTGCCA TTAATAAAAA 5340  
 AAAAAAAAAA AAAAAAAAAA AAAAAA

Seq ID NO: 575 Protein sequence:  
 Protein Accession #: Eos sequence

60  
 65  
 70  
 75  
 80  
 85

1 11 21 31 41 51  
 MRILKRFPLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALNQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQNVNVLKKL KFGQMDKTSI ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITPHW GKQNMSSDGS EHSLEQKFPF LEMQIYCFDA DRFSSFEEAV KGKGLRLALS 180  
 ILFEVGTEEN LDFKALIDGV ESVSRFGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMO QSGYVLMMDY LQNNFREQQY KFERQVFSSY 300  
 TGKEEIHFAV CSSEFENVQA DPNYVLSLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTDGVDQ LGAILNNLLP NMSYVLQIVA ICTNGLYGY SDQLIVDMET DNPELDLPPE 420  
 LIGTEEIIKE EEEKDIIEG AIVNPGRDSA TNQIRKKEPO ISTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDVPNTSL NSTSQPVTKL ATEKDISLTS QTVELPHT PHT VEGTASLND 540  
 GSKTVLRSFH MNLGTAESL NTVSITEYES ESLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGIIFS SENPETITYD VLIPEARNA SEDSTSSGSE ESLKDPMEG NVWFPSSTDI 660  
 TAQPDVSGSR ESFLQNTYTE IRVDESEKTT KSFGAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAPT SSRODLVST VNVVYSQTTQ PVYNAEASNS SHESRIGLAE GLESEKKAVI 780  
 PLVIVSALTP ICLVVLVGIL IYWRKCFQTA HFVLEDSTSP RVISTPPTPI FPISEDDVGA 840  
 PIKHFPHKVA DLHASSGFTE EPETLKEFYQ EVQSCVDLIG ITADSSNHFD NKHKRYINI 900  
 VAYDHSRVKL AQLAEKDGKL TDYINANYVD GYNRPKAYIA AQGPLKSTAE DFRWMIWERN 960  
 VEIVIMLTNL VEKGRRCQD YWADGSEET GNFLVTQKSV QVLAAYTVRN FTLRNTKIKK 1020  
 GSQGRPSGR VVTQYHYTQW PDMGVPEYSL PVLTFVRKAA YAKRHAVGPV VVHCAGVGR 1080  
 TGTIVLDSM LQQIQHBTG NIFGFLKHIR SORNYLVQTE EQVYFIHDTL VEALLSKETE 1140  
 VLDSHIAHYV WALLIPGPAG KTKLEKQFQL LSQSNIQQSD YSALKQCNR EKNRTSSIIP 1200  
 VERSRVGLIS LSGEGTDYIN ASYIMGYYS NEFIITQHPL LHTIKDFWRN IWDHNAQLVV 1260  
 MIPDQGNMAE DEFVYWNPKD EPINCESFKV TLMABEHKCL SNEEKLIQD FILEATQDDY 1320  
 VLEVRHFCPP KNPNDSPIS KTFELISVIK EBAANRDGPM IVDEHGGVT AGTFPCALTTL 1380  
 MHQLEKENSU DUYQVAKMIN LMRPGVFADI EQYQFLYKVI LSLVSTRQEE NPSTSLDSNG 1440  
 AALPDGNIAB SLESIV

Seq ID NO: 576 DNA sequence  
Nucleic Acid Accession #: EOS sequence  
Coding sequence: 148-4494

5	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCACTTCTGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
10	CGGCGAGGGG	CGCGAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGGCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTTGAAG	AGATTGGCTG	GTCTATACAC	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAACCTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
15	AACACATTCA	TTCAATACAC	TGGGAAACAA	GTGGAAATTA	ATCTCACTAA	TGACTACCGT	480
	GTACGCGGAG	GAGTTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCAGTGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCAGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAAAGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAAATTG	720
20	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TGATCTCTTT	GAACTTCTG	CCAACTCAA	CTGACAAGTA	TTACATTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTGTGTG	AAGTTCCTAC	AATGCAACAA	960
	TCTGTTATG	TGATCTGAT	GGACTACTTA	CAAAACAATT	TTGGAGAGCA	ACAGTACAAG	1020
25	TTCTCTAGAC	AGGTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTTCATGA	AGCAGTTTGT	1080
	AGTTGAGAAC	CAGAAATATG	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTTACA	1140
	TGGGAAAGAC	CTGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGTCTATT	TCAATAATTT	GCTACCCAAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
30	TGCATTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGAGCAT	GCCTACTGAT	1380
	AATCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAATCT	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATCA	GGAAAAAGGA	ACCCAGATT	TCTACCACAA	CACACTACAA	TGCGATAGGG	1560
35	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCCTTGGA	1620
	AAGGGTGATG	TTCCCAATAT	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTTCCTT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGTG	1740
	GAAGGTACTT	CAGCCTCTTT	AAATGATGGC	TCTAAAACCT	TTCTTAGATC	TCCACATATG	1800
	AACTGTGCGG	GGACTGCGAG	ATCCTTAAAT	ACAGTTTCTA	TAACAGATA	TGAGGAGGAG	1860
40	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTCTCT	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAAACCCAG	AGACATAAAC	ATATGATGTC	CTTATACCCG	AATCTGCTAG	AAATGCTTCC	2040
	GAAGATTCAA	CTTCACTCAG	TTCAAGAGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGGTTCT	CTAGCTCTCT	AGACATAACA	GCACAGCCCG	ATGTTGATC	AGGCAGAGAG	2160
45	AGCTTTCTCC	AGACTAATTA	CACTGAGATA	CGTGTGTATG	AATCTGAGAA	GACAACCAAG	2220
	TCCTTTCTGA	CAGGCCCACT	GATGTCAAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTTGCTTA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCAGACAAAC	AGGATTGGT	CTCCAAGGTC	AACTGGGTAT	ACTGCGAGAC	AACCCCAACG	2400
	GTATCAATAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
50	GAATCCGAGA	AGAAGGCCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTATCTGTT	2520
	CTAGTGGTTC	TTTGGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCCAACCC	TATCTTTCCA	2640
	ATTTGAGATG	ATGTCGAGC	AATTCCAATA	AAGCACTTTC	CAAGGATGTT	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAATTT	GAGGAAGTGC	AGAGCTGTAC	TGTTGACTTA	2760
55	GGTATTACAG	CAGACAGCTC	CAACCAACCA	GACAACAGC	ACAAGAATCG	ATACATAAAT	2820
	ATCGTTGCTT	ATGATCATAG	CAGGGTTAAG	CTAGCACAGC	TTGCTGAAAA	GGATGGCATA	2880
	CTGACTGATT	ATATCAATGC	CAATTATGTT	GATGGCTACA	ACAGACCAAA	AGCTTATATT	2940
	GCTGCCCAAG	GGCCACTGAA	ATCCACAGCT	GAAGATTCTT	GGAGAAATGAT	ATGGGAACAT	3000
	AAATGGAAG	TTATTGTCT	GATAACAAAC	CTCGTGGAGA	AAGGAAGGAG	AAATGTGAT	3060
	CAGTACTGGC	CTGCGGATGG	GAGTGAGGAG	TACGGGAAC	TTCTGTGCTC	TCAGAGAGAT	3120
60	GTGCAAGTGC	TTGCCATTAT	TACTGTGAGG	AATTTTACTC	TAAGAAACAC	AAAAATAAAA	3180
	AAGGGCTCCC	AGAAAGGAAG	ACCCAGTGGG	CGTGTGGTCA	CACAGTATCA	CTACAGCGAG	3240
	TGGCCTGACA	TGGGAGTACC	AGAGTACTCC	CTGCCAGTGC	TGACCTTTGT	GAGAAAGGCA	3300
	GCCTATGCCA	AGCGCCATGC	AGTGGGCCCT	TTGTGCTGTC	ACTGCACTGC	TGGAGTTGGA	3360
	AGAACAGGCA	CATATATTGT	GCTAGACAGT	ATGTTGCAGC	AGATTCAACA	CGAAGGAACT	3420
65	GTCAACATAT	TTGGCTTCTT	AAAACACATC	CGTTCAACAA	GAATTTATT	GGTACAAACT	3480
	GAGGAGCAAT	ATGCTTCTAT	TCATGATACA	CTGGTTGAGG	CCATACCTTAG	TAAAGAACT	3540
	GAGGTGCTGG	ACAGTCAAT	TCATGCCTAT	GTAAATGCAC	TCCTCATTTCC	TGGACCGACA	3600
	GGCAAAACAA	AGCTAGAGAA	ACAATTCCAG	CTCCTGAGCC	AGTCAAAAT	ACAGCAGAGT	3660
	GACTATTCTG	CAGCCCTAAA	GCAATGCAAC	AGGGAAGAGA	ATCGAACTTC	TTCTATCATC	3720
70	CCTGTGAAAA	GATCAAGGGT	TGGCATTTCA	TCCCTGAGTG	GAGAAGGCAC	AGACTACATC	3780
	AATGCCCTCT	ATATCATGGG	CTATTACCAG	AGCAATGAAT	TCATCAATTAC	CCAGCACCTT	3840
	CTCCTTCTATA	CCATCAAGGA	TTTCTGGAGG	ATGATATGGG	ACCATTAATGC	CCAAGTGGTG	3900
	GTATATGATT	CTGATGGCCA	AAACATGGCA	GAAGATGAAT	TTGTTTACTG	GCCAAATAAA	3960
	GATGAGCCTA	TAAATTGTGA	GAGCTTTAAG	GTCACTCTTA	TGGCTGAAGA	ACACAAATGT	4020
75	CTATCTAATG	AGGAAAAACT	TATAATTACG	GACTTTATCT	TAGAAGCTAC	ACAGGATGAT	4080
	TATGTACTTG	AAGTGAGGCA	CTTTCAGTGT	CTAAATGGGC	CAAAATCCAGA	TAGCCCACTT	4140
	AGTAAACCTT	TTGAACCTAT	AAGTGTATTA	AAAGAAGAAG	CTGCCAATAG	GGATGGGCCCT	4200
	ATGATTGTTT	CTGATGAGCA	TGGAGGAGTG	ACGGCAGGAA	CTTTCTGTGC	TCTGACAACC	4260
	CTTATGCAAC	AACTAGAAAA	AGAAAAATTC	GTGGATGTTT	ACCAAGGTAGC	CAAGATGATC	4320
80	AACTGTGATG	GGCCAGGAGT	CTTTGCTGAC	ATTGAGCAGT	ATCAGTTTCT	CTACAAAGTG	4380
	ATCCTCAGCC	TTGTGAGCAC	AAGGCAGGAA	GAGAATCCAT	CCACCTCTCT	GGACAGTAAT	4440
	GGTGAGCAT	TGCCGTGATG	AAATATAGCT	GAGAGCTTAG	AGTCTTTAGT	TTAACACAGA	4500
	AAGGGGTGGG	GGGACTCACA	TCTGAGCATT	GTTTTCTCT	TCCTAAAAAT	AGGCAGGAAA	4560
	ATCAGTCTAG	TTCTGTTATC	TGTTGATTTT	CCATCACCTG	ACAGTAACCT	TCATGACATA	4620
85	GGATTCTGCC	GCCAAATTTA	TATCATTAAAC	AAATGTGTC	TTTTTGCAAG	ACTTGTAAAT	4680
	TACTTATTAT	GTTTGAACCTA	AAATGATTGA	ATTTTACAGT	ATTTCTAAGA	ATGGAATTGT	4740
	GGTATTTTTT	TCTGTATTGA	TTTTAACAGA	AAATTTCAAT	TTATAGAGGT	TAGGAATTCC	4800

AAACTACAGA AATGTTTGT TTTTAGTGTC AAATTTTAG CTGTATTGT AGCAATTATC 4860  
 AGGTTTGCTA AATATATAAC TTTTAATACA GTAGCCTGTA AATAAAACAC TCTTCCATAT 4920  
 GATATTCAAC ATTTTACAAC TGCAGTATTC ACCTAAAGTA GAAATAATCT GTTACTTATT 4980  
 GTAAATACTG CCTAGTGTGC TCCATGGACC AAATTTATAT TTATAATTGT AGATTTTAT 5040  
 ATTTTACTAG TGAGTCAAGT TTTCTAGTTC TGTGTAATGT TTAGATTAA TGACGTAGTT 5100  
 CATTAGCTGG TCTTACTCTA CCAATTTTCT GACATTGTAT TGTGTTACCT AAGTCATTA 5160  
 CTTTGTTCAC GCATGTAATT TTAACTTTGT TGGAAATAG AAATACCTTC ATTTTGAAAG 5220  
 AAGTTTTTAT GAGAATAACA CCTTACCAA CATTTGTCAT ATGGTTTTTA TCCAAGGAAT 5280  
 TGCAAAAATA AATATAATAA TTGCCATTAA AAAAAAATA AAAAAAATA AAAAAAATA 5340  
 AAA

Seq ID NO: 577 Protein sequence:  
 Protein Accession #: EOS sequence

1 11 21 31 41 51  
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQNVNMLKKL KFGQWMDKSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITPHW GKCNMSSDGS EHSLEGQKFP LEMQIYCFDA DRPSSFEEAV KGKGLRLALS 180  
 ILFEVGTSEN LDFKALIDGV ESVSFPGKQA ALDPFILLNL LPNSTDKYII YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLIMQ QSGYVLMNDY LQNNFREQQY KFSRQVFSY 300  
 TGKEBIHEAV CSSEPEVQA DPENYTSLLV TWERPRVVDY TMIEKPAVLY QQLDGEDQTK 360  
 HEFLTDGQD LGAILNLLP NMSYVLQIVA ICTNGLYGY SDQLIVDMPT DNFELDLFPE 420  
 LGTEBIKE EEEGRDIEEG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDVNTSL NSTSQFVTKL ATEKDISLTS QTVTELPHT VEGTSASLND 540  
 GSKTVLRSPH MNLSTGAEEL NTVSITTEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETIITYD VLIPESARNA SEDSTSSGSE ESLKDPMEG NVWFPSSTDI 660  
 TAQPDVSGSR ESPLQNTYE IRVDESEKTT KSFSAGPVM S QGPSVTDLEM PHYSTFAYFP 720  
 TEVTHAFTP SSRQDLVST VNVVYSQTTQ PVYNEASNS HESRIGLAEG LESEKKAVIP 780  
 LVIVSALTFI CLVVLVGLLI YWRKCFQTAH FYLEDSTSPR VISTPTPIF PISDDVGAIP 840  
 IKHPPKHVAD LHASGTFEEL FEEVQSTVD LGITADSSNH PDNKHKNRYI NIVAYDHSRV 900  
 KLAQALAEKD KLTDTYINANY VDGYNRPKAY IAAQGLKST AEDFWRMINE HNVFVIMIT 960  
 NLVEKGRKRC DQWVPADGSE EYGNFLVTQK SVQVLAITYV RNFTLRNTKI KKGSKGRPS 1020  
 GRVVTQYHYT QWPDGVPYEL SLPLVTFVRK AAYAKRHAVG PVVVCASAGV GRTGYIVLD 1080  
 SMLQIQIEHG TVNIPGLKXH IRSQRNVLVQ TEEQYVFIHD TLVEALLSKE TEVLDSHIEA 1140  
 YVMALLIPGP AGHTKLEKQF QLLSQSNLQO SDYSAALKQC NREKNTSSI IPVERSVGI 1200  
 SLSGEGEDY INASYIMGY QSNFPIITQH PLLETKIDFW RMIWDHNAQL VVMIPDQNM 1260  
 AEDEPVYVFN KDEPINCSEF KVTLMABEHK CLSNEEKLI QDFILEATQD DYLVEVREFQ 1320  
 CPKNPNDPSP ISKTPELISV IKEEAANRDG PMIVHDEHGG VTAGTFCAL TLMHOLEKEN 1380  
 SVDVYQVAKM INLMRPVFA DIEQYQFLYK VILSLVSTRO EENPSTSLDS NGAALPDGNI 1440  
 AESLESLV

Seq ID NO: 578 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 501-4514

1 11 21 31 41 51  
 CACACATACG CACGCACGAT CTCACCTTGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAAGAC ATTTCTCTCG CTCCTCTTCC CTCTCCACTC TGAGAGCAG AGGAGCGGCA 120  
 CGGCGAGGGG CGCAGACCGG TCTGGAATG CGAATCTCAA AGCGTTTCTT CGCTTGCAAT 180  
 CAGCTCTCTC GTGTTTGGCG CCTGGATTGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTGTGTAAG AGATTGGCTG GTCTATACA GGAGCACTGA ATCAAAAAAT TGGGGAAGA 300  
 AATATCCAA CAGTAAATAG CCAAAACAAT CTCCTATCAA TATTGATGAA GATCTTACAC 360  
 AAGTAAATGT GAATCTTAAG AAACCTAAAT TTCAGGGTTG GGATAAAACA TCATTGGAAA 420  
 ACACATTCAT TCATAACACT GGGAAACAG TGGAAATTA TCTCACTAAT GACTACCGTG 480  
 TCAGCGGAGG AGTTTCAGAA ATGGTGTTTA AAGCAAGCAA GATAACTTTT CACTGGGGAA 540  
 AATGCAATAT GTCACTGTAT GGATCAGAGC ATAGTTTAGA AGGACAAAAA TTTCCACTTG 600  
 AGATGCAAT TACTGCTTT GATGCGACC GATTTCAGG TTTTGAGGAA GCAGTCAAAG 660  
 GAAAAGGGAA GTTAAGAGCT TTATCCATT TGTTTGAGGT TGGGACAGAA GAAAATTGG 720  
 ATTTCAAGC GATTATGTAT GGAGTCGAAA GTGTTAGTGG TTTTGGGAG CAGGCTGCTT 780  
 TAGATCCATT CATACTGTGT AACCTTCTGC CAAACTCAAC TGACAAGTAT TACATTTACA 840  
 ATGGCTCATT GACATCTCCT CCTGCAACAG ACACAGTTGA CTGGATTGTT TTTAAGATA 900  
 CAGTTAGCAT CTCTGAAAGC CAGTTGGCTG TTTTGTGTA AGTCTTACA ATGCAACAAT 960  
 CTGGTTATGT CATGCTGATG GACTACTTAC AAAACAATT TCGAGAGCAA CAGTACAAGT 1020  
 TCTCTAGACA GGTGTTTTC TCATACACTG GAAAGGAAGA GATTTCATGA GCAGTTTGTA 1080  
 GTTCAGAAC AGAAAATGT CAGGCTGACC CAGAGAATTA TACCAGCCTT CTGTGTACAT 1140  
 GGGAAAGACC TCGAGTCGTT TATGATACCA TGATTGAGAA GTTTGCAGTT TTGTACCAGC 1200  
 AGTTGGATGG AGAGGACCAA ACCAAGCATG AATTTTGGAC AGATGGCTAT CAAGACTTGG 1260  
 GTGCTATTCT CAATAATTG CTACCCAATA TGAGTTATGT TCTTCAGATA GTAGCCATAT 1320  
 GCCTAATGG CTTATATGGA AAATACAGCG ACCAAGTAT TGTGACATG CCTACTGATA 1380  
 ATCTGAAC TGAATCTTCT CCTGAATTAA TTGGAACCTG AGAATAATC AAGGAGGAGG 1440  
 AAGAGGGAAA AGACATTGAA GAAGGCGCTA TTGTGAATCC TGGTAGAGAC AGTGCTACAA 1500  
 ACCAAATCAG GAAAAGGAAA CCCCAGATTT CTACCAACAC AACTACAAT CGCATAGGGA 1560  
 CGAAATACAA TGAAGCCAA ACTAACGAT CCCCACAAG AGGAAGTGAA TTCTCTGGAA 1620  
 AGGGTGATGT TCCCAATACA TCTTAAAT CCACCTCCCA ACCAGTCACT AAATTAGCCA 1680  
 CAGAAAAGA TATTTCTCTG ACTTCTCAGA CTGTGACTGA ACTGCCACT CACACTGTGG 1740  
 AAGGTACTTC AGCTCTTTA AATGATGGCT CTAAACATGT TCTTAGATCT CCACATATGA 1800  
 ACTTGTGCGG AGAGTCAGAA TCCTTAATA CAGTTTCTAT AACAGAATAT GAGGAGGAGA 1860  
 GTTTATGAC CAGTTTCAAG CTGTACTG GAGCTGAAGA TTCTTCAGGC TCCAGTCCG 1920  
 CAACCTCTGC TATCCATTC ATCTCTGAGA ACATATCCCA AGGGTATATA TTTTCTCCG 1980  
 AAAACCCAGA GACAATAACA TATGATGTCC TTATACCAGA ATCTGCTAGA AATGCTCCG 2040  
 AAGATTCAAC TTATCAGGT TCAGAAAGAT CACTAAAGGA TCCTTCTATG GAGGGAATG 2100  
 TGTGTTTCC TAGCTCTACA GACATAACAG CACAGCCCGA TGTGATCA GGCAGAGAGA 2160  
 GCTTCTCCA GACTAATTAC ACTGAGATAC GTGTTGATGA ATCTGAGAG ACAACCAAGT 2220  
 CCTTTCTGC AGGCCACGTG ATGTACAGG GTCCCTCAGT TACAGATCTG GAAATGCCAC 2280

ATTATTCTAC CTTTGCCTAC TTCCCACTG AGGTAACACC TCATGCTTTT ACCCCATCCT 2340  
 CCAGACAACA GGATTGGTTC TCCAGGTC ACGTGGTATA CTCGAGACA ACCCAACC3G 2400  
 TATACAATGA GGCCAGTAAT AGTAGCCATG AGTCTCGTAT TGGTCTAGCT GAGGGGTGG 2460  
 AATCCGAGAA GAAGGCGAGTT ATACCCCTTG TGATCGTGT AGCCCTGACT TTTATCTGTC 2520  
 TAGTGGTCTT TGTGGGTATT CTCACTACT GGAGGAAATG CTTCCAGACT GCACACTTTT 2580  
 ACTTAGAGGA CAGTACATCC CCTAGAGTTA TATCCACACC TCCAACACCT ATCTTTCCAA 2640  
 TTTAGATGA TGTGCGAGCA ATTCCAATAA AGCACTTTCC AAAGCATGTT GCAGATTAC 2700  
 ATGCAAGTAG TGGGTTTACT GAAGAATTG AGACACTGAA AGAGTTTAC CAGGAAGTGC 2760  
 AGAGCTGTAC TGTGACTTA GGTATTACAG CAGACAGCTC CAACCAACCA GACAACAAGC 2820  
 ACAAGAAATG ATACATAAAT ATCGTTGCCT ATGATCATAG CAGGGTTAAG CTAGCACAGC 2880  
 TTGCTGAAAA GGATGGCAAA CTGACTGATT ATATCAATGC CAATTATGTT GATGGCTACA 2940  
 ACAGACCAAA AGCTTATATT GCTGCCAAG GCCCACTGAA ATCCACAGCT GAAGATTCT 3000  
 GGAGAAATGAT ATGGGAACAT AATGTGGAAG TTATTGTCAT GATAACAAAC CTCGTGGAGA 3060  
 AAGGAAGGAG AAAATGTGAT CAGTACTGGC CTGCCGATGG GAGTGAGGAG TACGGGAAC 3120  
 TTCTGGTCTC TCAGAAAGAT GTGCAAGTGC TTGCCATTAT TACTGTGAGG AATTTTACTC 3180  
 TAAGAAACAC AAAAATAAAT AAGGGCTCCC AGAAAGGAAG ACCCAGTGGG CGTGTGGTCA 3240  
 CACAGTATCA CTACACGCGAG TGGCCTGACA TGGAGTATCC AGAGTACTCC CTGCCAGTGC 3300  
 TGACCTTTGT GAGAAAGGCA GCCTATGCCA AGCGCCATGC AGTGGGGCCT GTTGTCTGCC 3360  
 ACTGCAGTGC TGGTGTGGA AGAACAGGCA CATATATTGT GCTAGACAGT ATGTTGAGC 3420  
 AGATTCAACA CGAAGGAAGT GTCAACATAT TTGGCTTCTT AAAACACATC CGTTCACAA 3480  
 GAAATATTAT TATCAAAAT GAGGAGCAAT ATGCTTTCAT TCATGATACA CTGGTTGAGG 3540  
 CCATCTTAG TAAAGAAACT GAGGTGCTGG ACAGTCATAT TCATGCCTAT GTTAATGCAC 3600  
 TCCTCATTC TGGACCGACA GGCAAAACAA AGCTAGAGAA ACAATTCAG CTCCTGAGCC 3660  
 AGTCAAAATAT ACAGCAGAGT GACTATTCTG CAGCCCTAAA GCAATGCAAC AGGGAAGA 3720  
 ATCGAACTTC TTCTATCATC CCTGTGAAA GATCAAGGGT TGGCATTTC TCCCTGAGTG 3780  
 GAGAAAGGCA AGACTACATC AATGCCCTCT ATATCATGGG CTATTACAG AGCAATGAAT 3840  
 TCATCATTAC CCAGCACCTC CTCTTCATA CCATCAAGGA TTTCTGGAGG ATGATATGGG 3900  
 ACCATAATGC CCAACTGGTG GTTATGATTC CTGATGGCCA AAACATGGCA GAAGATGAAT 3960  
 TTGTTTACTG GCCAAATAAA GATGAGCCTA TAAATTGTGA GAGCTTAAAG GTCACCTTA 4020  
 TGGCTGAAGA ACACAAATGT CTATCTAATG AGGAAAACT TATAATTAC GACTTTATCT 4080  
 TAGAAGCTAC ATACAGATGAT TATGTACTTG AAGTGAGGCA CTTTCAGTGT CCTAAATGGC 4140  
 CAAATCCAGA TAGCCCCATT AGTAAACTT TTGAACCTAT AAGTGTATA AAAGAAGAAG 4200  
 CTGCCAATAG GGTGGGCTT ATGATTGTTT ATGATGAGCA TGGAGGAGTG ACGGCAGGAA 4260  
 CTTTCTGTGC TCTGCAAAAC CTTATGCACC AACTAGAAAA AGAAAAATCC GTGGATGTTT 4320  
 ACCAGGTAGC CAAGATGATC AATCTGATGA GGCCAGGAGT CTTTGTCTGAC ATTGAGCAGT 4380  
 ATCAGTTTCT ATCAAAAGT ATCCCTCAGCC TTGTGAGCAC AAGGCAGGAA GAGATCCAT 4440  
 CCACCTCTCT GGACAGTAAT GGTGCAGCAT TGCCGTAGTG AAATATAGCT GAGAGCTTAG 4500  
 AGTCTTTAGT TTAACACAGA AAGGGGTGGG GGGACTCACA TCTGAGCATT GTTTCTCTCT 4560  
 TCCTAAAATT AGGCAGGAAA ATCAGTCTAG TTCTGTTATC TGTGATTTT CCATCACCTG 4620  
 ACAGTAATCT TCTAGACATA GGATTCTGCC GCCAAATTA TATCATTAA ATGTGTGCC 4680  
 TTTTGTGAAG ATCGTAAAT TACTATTAT GTTTGAACCT AAATGATTGA ATTTACAGT 4740  
 ATTTCTAAGA ATGGAATGTG GGTATTTTT TCTGTATTGA TTTTAACAGA AAATTTCAAT 4800  
 TATAGAGGT TAGGAATTC AACTACAGA AAATGTTGT TTTTAGTCTC AAATTTTAG 4860  
 CTGTATTGT AGCAATTATC AGGTTTGCTA GAAATATAAC TTTTAATACA GTAGCCTGTA 4920  
 AATAAAACAC GTTCTCATAT GATATTCAAC ATTTTACAAC TGCAATTATC ACCTAAAGTA 4980  
 GAAATAATCT CTTCTTATG GTAAATAGT CCCTAGTGTC TCCATGGACC AAATTTATAT 5040  
 TTATAATGT AGATTTTAT ATTTTACTAC TGAGTCAAGT TTTCTAGTTC TGTGTAATG 5100  
 TTTAGTTTAA TGAGCTAGT CATTAGCTGG TCTTACTCTA CCAGTTTCT GACATGTAT 5160  
 TGTGTACCT AAGTCATTAA CTTTGTTTCA GCATGTAATT TTAACTTTG TGGAAATAG 5220  
 AATACCTTC ATTTTGAAG AAGTTTTAT GAGATAACA CCTTACCAA CATTTCTCA 5280  
 ATGTTTTTA TCCAAGGAA TGCAAAATA AATATAATA TTGCCATTAA AAAAAAAA 5340  
 AAAAAAAA AAA

Seq ID NO: 579 Protein sequence:  
 Protein Accession #: EOS sequence

1 11 21 31 41 51  
 MYPKASKITP HWGKCNMSSD GSEHSLEGQK PPLEMQIYCF DADRFSSPFE AVKQKGLRA 60  
 LSILFEVGT EMLDFKAIID GVESVSRFGK QAALDPFILL NLLPNSTDKY YIYNGSLTSP 120  
 PCTDVTWIV FKTIVSISES QLAVFCEVLT MQQSGYVMLM DYLNQNFREQ QYKFSRQVFS 180  
 SYTGKEIHE AVCSSEPENV QADPENYTSI LVTWERPRV YDTMIEKFAV LYQLDGEDQ 240  
 TKHEFLTGY QDLGAILNML LPNMSYVLQI VAICTNGLYG KYSDQLIVDM PTDNFELDLE 300  
 PELIGTEII KEEEBGKDIE EGAIVNPGRD SATNQIRKKE PQISTTHYIN RIGTKYNEAK 360  
 NRSPTRGSE FSGKDVPT SLNSTSQPV KLA TEKDISL TSQTVTELP HTVEGTBASL 420  
 NDGSKTVLRS PEMNLGTAE SLNTVSI TEY EESLLTSFK LDTGAEDSSG SSPATSAPF 480  
 ISENISQGYI FSSSEPTIT YDVLIPESAR NASEDSTSSG SEESLKOPSM EGNVWFPSST 540  
 DITAPQDVGS GRESFLQNTY TEIRVDESEK TTKFSAGPV MSQGPVTDL EMPHYSTPAY 600  
 FPTEVTPHAP TPSSRQQLV STVNVSYSQT TQPVYNEASN SSHESTRIGLA EGLESEKKAV 660  
 IPLVIVSALT FICLVVLVGI LIYWRKCFQT AHFYLEDSTS PRVISTPPTP IFPIISDDVGA 720  
 IPIKHPPKHV ADLHASSGFT EEPETLKEFY QEVQSCVTDL GITADSSNHP DNKHKNRYIN 780  
 IVAYDHSRVK LAQLAEKDGK LTDYINANYV DGYNRPKAYI AAQGPLKSTA EDFWRMIWEH 840  
 NVEVIMVITN LVEKGRKCD QYWPADGSEB YGNFLVTQKS QVFLAYYTVR NPTLRNTKIK 900  
 KGSQGRPSG RVVTQYHYTQ WPDGVPPEYS LPVLTFRKA AYAKRHAVGP VVVHCSAGVG 960  
 RTGTVIVLDS MLQIQHEGT VNIFGFLKHI RSQRNYLVQT BEQVVFIDHT LVEAILSSET 1020  
 EVLDSHIIAY VNALLIPQPA GKTKLEKQFO LLSQSNIQQS DYSALKQCN REKNRTSSII 1080  
 FVERSRVIGS SLSEBGTDYI NASYIMGYQ SNEFIITQHP LHRTIKDFWR MIWDENAQLV 1140  
 VMIPDQNMMA EDEFVYWFNK DEPINCSEFK VTLMAEHEKC LSNREKLIQ DPFILEATQDD 1200  
 YVLEVRHFQC PKNPNPSPSI SKTFELISVI KEEAANDGCP MIVHDEHGGV TAGTFCALTT 1260  
 LMHOLEKENS VDVYQVAKMI NLMRPGVFAD IEQYQFLYKV ILSLVSTRQE ENPSTSLDSN 1320  
 GAALPDGNIA ESLESV

Seq ID NO: 580 DNA sequence  
 Nucleic Acid Accession #: EOS sequence  
 Coding sequence: 148-4632

1 11 21 31 41 51



	CACACATAAG	CACGCACGAT	CTCACTTGG	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAAC	ATTTCCTTCG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
5	CGCGAGGGG	CGCGAGACCG	TCTGGAAATG	CGAATCCTAA	AACGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTGGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
	CTTGTGGAAG	AGATTGGCTG	GTCCATATCA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CATGTAATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
10	AACACATTCA	TTCAATACAC	TGGGAAAAAC	GTGGAATTA	ATCTCACTAA	TGACTACCGT	480
	GTACAGCGAG	GAGTTTCAGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCATCGGGGA	540
	AAATGCAATA	TGTCATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCCAA	TCTACTGCTT	TGATGCCGAC	CGATTTTCAA	GTTTTGAGGA	AGCAGTCAAA	660
	GGAAAGGGG	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTGG	720
15	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTAGTC	GTTTTGGGAA	GCAGGCTGCT	780
	TTAGATCCAT	TCATCTGTTT	GAACTTCTG	CCAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTTAAGAT	900
	ACAGTTAGCA	TCTCTGAAG	CCAGTTGGCT	GTTTTTTTGT	AAGTCTTAC	AATGCAACAA	960
	TCTGTTATG	TGATGCTGAT	GGACTACTTA	CAAAACAATT	TTGAGAGCA	ACAGTACAAG	1020
20	TTCTCTAGAC	AGCTGTTTTC	CTCATACACT	GGAAAGGAAG	AGATTCACTA	AGCAGTTTGT	1080
	AGTTTCAAGC	CAGAAATGTT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTTGTATCA	1140
	TGGGAAAGAC	CTCGAGTCGT	TTATGATACC	ATGATTGAGA	AGTTTGCAAT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAGAGCTTG	1260
	GGTGCTATTG	TCAATTAATT	GCTACCCAA	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
25	TGCACTAATG	CGTTATATGG	AAAATACAGC	GACCAACTGA	TTGTCGACAT	CGCTACTGAT	1380
	AATCTCTAAG	TTGATCTTTT	CCCTGAATTA	ATGGAACTG	AAGAAATAAT	CAAGGAGGAG	1440
	GAAGAGGGAA	AGACATTGAT	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAATATCA	GGAAAGGAAG	ACCCAGGATT	TCTACCACAA	CACACTACAA	TCGCATAGGG	1560
	ACGAATATCA	ATGAAGCCAA	GACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
30	AAGGTTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAAG	ATATTCTCTT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGTG	1740
	GAAGGTACTT	CAGCTCTCTT	AAATGATGGC	TCTAAACTG	TTCTTAGATC	TCCCATATG	1800
	AACCTTGTGG	GGAATGACGA	ATCCTTAAAT	ACAGTTTCTA	TAAACAGATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
35	GCAACTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGTATAT	ATTTTCTCTC	1980
	GAAAAACCA	AGACATATAC	ATATGATGTC	CTTATACCA	AATCTGTCTG	AAATGCTTCC	2040
	GAAGATTCAA	TTCTCATCAG	TTCAAGAAAG	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCTCC	AGACTAATTA	CACCTGAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
40	TCCTTTCTG	CAGGCCAGT	GATGTCACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATTATTCTA	CTTTGGCCTA	CTTCCCAACT	GAGGTAAAC	CTCATGCTTT	TACCCCATCC	2340
	TCAGACAAAC	AGGATTGCTG	CTCCACGGTC	AACGTGGTAT	ACTGCAGAC	AACCCCAACG	2400
	GTATACAAAT	AGGCCAGTAA	TAGTAGCCAT	GAGTCTCGTA	TTGGTCTAGC	TGAGGGGTTG	2460
	GAATCCGAGA	AGAAGGCAGT	TATACCCCTT	GTGATCGTGT	CAGCCCTGAC	TTTTATCTGT	2520
45	CTAGTGGTTC	TTGTGGGTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACACTTT	2580
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCAACACC	TATCTTTCCA	2640
	ATTTAGATG	ATGTCGAGC	AATTTCAATA	AAGCACTTTC	CAAGCATGTG	TGCAGATTTA	2700
	CATGCAAGTA	GTGGGTTTAC	TGAAGAAATT	GAGACACTGA	AAGAGTTTAA	CCAGGAAGTG	2760
	CAGAGCTGTA	CTGTTGATCT	AGGTATTACA	GCAGACAGCT	CCAACCAACC	AGACAACCAAG	2820
50	CACAAGAAAT	GATACATAAA	TATCGTTGCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	2880
	CTTGCTGAAA	AGGATGGCAA	ACTGACTGAT	TATATCAATG	CCAATTATGT	TGATGGCTAC	2940
	AACAGACCAA	AGCTTTATAT	TGCTGCCCAA	GGCCCACTGA	AATCCACAGC	TGAAGATTTC	3000
	TGGAGAAATG	TATGGGAACA	TAAATGTGAA	GTATTGTGTA	TGATAACAAA	CCTCGTGGAG	3060
	AAAGGAAGGA	GAAATGTGTA	TCAGTACTGG	CCTGCCGATG	GGAGTGAGGA	GTAACGGAAAC	3120
55	TTTCTGGTCA	CTCAGAAGAG	TGTGCAAGTG	CTTGCTTATT	ATCTGTGAG	GAATTTTACT	3180
	CTAAGAAACA	CAAAAAATA	AAAGGGCTCC	CAGAAAGGAA	GACCCAGTGG	ACGTGTGGTC	3240
	ACACAGTATC	ACTACACGCA	GTGGCCTGAC	ATGGGAGTAC	CAGAGTACTC	CCTGCCAGTG	3300
	CTGACCTTTG	TGAGAAAGGC	AGCCTATGCC	AAGCGCCATG	CAGTGGGGCC	TGTTGTCTCT	3360
	CACCTGAGTG	CTGGAGTTGG	AAGAACAGGC	ACATATATTT	TGCTAGACAG	TATGTTGTCAG	3420
60	CAGATTCAAC	ACGAAGGAAC	TGTCAACATA	TTTGGCTTCT	TAAACACAT	CCGTTCCACA	3480
	AGAAATATT	TGTTACAAAC	TGAGGAGCAA	TATGTTCTCA	TTATGATAC	ACTGGTTGAG	3540
	GCCATACTTA	GTAAAGAAAC	TGAGGTGCTG	GACAGTCATA	TTATGCTCTA	TGTTAATGCA	3600
	CTCCTCATTC	CTGGACACGC	AGGCAAAACA	AAGCTAGAGA	AACATTTCCA	GGGTCTCACT	3660
	CTGTCAACCA	GGCTGGAGTG	CAGAGGCACA	ATCTCGGCTC	ACTGCAACCT	TCTCTCTCCT	3720
	GGCTTAACCT	ATCTCTCTAC	CTCAGCCTCC	CGATGGGCTG	GGACTATACT	CCTGAGCCAG	3780
65	TCAATATATC	AGCAGAGTGA	CTATTCTGCA	GCCCTAAAGC	AATGCAACAG	GGAAAGAAAT	3840
	CGAATCTCTT	CTATCATCCC	TGTGGAAAGA	TCAAGGGTTG	GCATTTATC	CCTGAGTGGG	3900
	GAAGGCACAG	ACTACATCAA	TGCTCTCTAT	ATCATGGGCT	ATTACCAAGG	CAATGAATTC	3960
	ATCATTACCC	AGCACCTCTT	CCTTCATACC	ATCAAGGATT	CTGAGGAGAT	GATATGGGAC	4020
70	CATAATGCC	AACTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	4080
	GTTTACTGGC	CAAAATAAGA	TGAGCCTATA	AATTGTGAGA	GCTTTAAGGT	CACCTCTATG	4140
	GCTGAAGAAC	ACAAATGTCT	ATCTAATGAG	GAAAACTTA	TAAATCAGGA	CTTTATCTTA	4200
	GAAGCTACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCACT	TTCAATGTCC	TAAATGGCCA	4260
	AATCCAGATA	GCCCCATTAG	TAAACCTTTT	GAACTTATAA	GTGTTATAAA	AGAAGAAGCT	4320
75	GCCAAATAGG	ATGGGCTCAT	GATTTGTCAT	GATGAGCATG	GAGGAGTGAC	GGCAGGAATC	4380
	TTCTGTGCTC	TGACAAACCT	TATGCAACCA	CTAGAAAAAG	AAAAATTCOG	GGATGTTTAC	4440
	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TTGCTGACAT	TGAGCAGTAT	4500
	CAGTTTCTCT	ACAAAGTAT	CCTCAGCCTT	GTGGGCACAA	GGCAGGAAGA	GAATCCATCC	4560
	ACCTCTCTGG	ACAGTAAATG	TGCAGCATTG	CCTGATGGAA	ATATAGCTGA	GAGCTTAGAG	4620
80	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	4680
	CTAAATATTG	GAGGAAATAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCG	ATCACTTGAC	4740
	AGTAACTTTC	ATGACATAGG	ATTTCTGGCG	CAAAATTATA	TCAATTAACA	TGTGTGCTTT	4800
	TTTGAAGAC	TTGTAATTTA	CTTATTATGT	TTGAACATAA	ATGATTGAAT	TTTACAGTAT	4860
	TTCTAAGAA	GGAATTTGGG	TATTTTTTTC	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	4920
	ATAGAGGTTA	GGAATTTCAA	ACTACAGAAA	ATGTTTGTTC	TTAGTGTCAA	ATTTTGTAGT	4980
85	GATTTGTAG	CAATTTACAG	GTTTGTCTAG	AATATAACTT	TTAATACAGT	AGCCTGTAAA	5040
	TAAACACTCT	TTCCATATGA	TATTCACAT	TTTACAACCT	CAGTATTAC	CTAAAGTAGA	5100
	AATAATCTGT	TACTTATTGT	AAATACTGCC	CTAGTGTCTC	CATGACCAAA	ATTTATATTT	5160



ATAATGTAG ATTTTATAT TTTACTACTG AGTCAAGTTT TCTAGTCTG TGTAAITGTT 5220  
 TAGTTTATG ACGTAGTTCA TTAGCTGGTC TTACTCTACC AGTTTCTGA CATGTATTG 5280  
 TGTACTCTAA GTCAATTAAT TTGTTTCAGC ATGTAAITTT AACTTTGTG GAAATAGAA 5340  
 ATACCTTCAT TTGAAAGAA GTTTTATGA GAATAACACC TTACCAACA TTGTTCAAT 5400  
 5 GGTTTTATC CAAGGAATTG CAAAAATAA TATAATATT GCCATTAAAA AAAAAA 5460  
 AAAAAA AAAAAA A

Seq ID NO: 581 Protein sequence:  
 Protein Accession #: EOS sequence

10 1 11 21 31 41 51  
 | | | | |  
 MRILKRFLAC IQLLCVCRLD WANGYYRQQR KLVEEIGWSY TGALEKQKNG KYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKKL KFGQWDTSL ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 15 PKASKITPHW GKCNSSSGS EHSLEGQKFP LEMQIYCFDA DRFSSEFVAV KKGKGLRALS 180  
 ILFEVGTEN LDFKAIIDGV ESVSFPGKA ALDPFILLNL LPNSTDKYII YNGSLTSPPC 240  
 TDTVDWIVFK DTVSISESQL AVFCEVLTMQ QSGYVLMNDY LQNNFREQQY KPSRQVFSSY 300  
 TGKEEHEAV CSSEPEVQA DPENYTSLLV TWERPRVVDY TMIEKPAVLY QQLDGEDQTK 360  
 20 HEFLTDGVOD LGAILNLLP NMSYVLQIVA ICTNGLYGYK SDQLIVDMPT DNPFLDLFPE 420  
 LIGTEBIKE EEEGKDIEEG AIVNPGRDSA TNQIRKKEPO ISTTHYNNRI GTKYNEAKTN 480  
 RSPTRGSEFS GKGDVNTSL NSTSQPVTKL ATEKDISLTS QVTTELPPHT VEGTSASLND 540  
 GSKTVLRSPH MNLGTAESL NTVSITEYSE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS SENPETIYD VLIPEASRNA SEDSTSSGSE ESLKDPSEMG NWFFPSSTDI 660  
 25 TAQPDVGSGR ESFLQNTYNE IRVDESEKTT KSFSAGPVMG QGPSVTDLEM PHYSTFAYFP 720  
 TEVTHAFTPT SSRQDLVST VNVVYSQTQ PVYNEASNSS HESRIGLAEG LESEKKAVIP 780  
 LVIVSALTPI CLVVLVGILI YWRKCFQTAH FYLEDSTSPR VISTPPTPIF PISDDVGAIP 840  
 IKHFPKHVAD LHASSGFTEE FETLKEFYQE VQSCTVDLGI TADSNHNPON KHNRYINIV 900  
 AYDHSRVLKA QLAEKDGKLT DYINANYVDG YNRPKAYIAA QGPLKSTAEF FWRMIWEHNV 960  
 30 EVIVMTNLV EKRERKCDQY WPADGSEYFG NPLVTQKSVQ VLAYYTVRNP TLRNKTKKG 1020  
 SOKGRPSGRV VTQHYHTQWP DMGVPEYSLP VLTFRVKAAY AKRHAVGPVV VHCAGVGR 1080  
 GTYIVLSDMS QQIQHEGTVN IPGFLKHRS QRNYLVQTEE QYVPIHDLTV EAILSKEDEV 1140  
 LDSEIHAYVN ALLIPGPAGK TKLEKQFQGL TLPRLCECRG TISAHCNLP LGLTDPPTSA 1200  
 SRVAGTILLS QSNIQSDYS AALKQCNREK NRTSIIIPVE RSRVGISSLS GBGTDYINAS 1260  
 35 YIMGYQYNE FIITQHPLLH TIKDFWRMIW DHNAQLVVMV PDGQMAEDE FYWPNKDEP 1320  
 INCESFKYTL NAEHKKCLSN EEKLIQDFI LEATQDDYVL EVRHFQCPKW PNPDSPISTK 1380  
 FELISVKEE AENRGGPMIV HDEHGGVTAG TFCALTLMH QLEKENSVDV YQVARMINLM 1440  
 RGPVPADIEQ YQFLYKVLIS LVGTRQENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 582 DNA sequence  
 Nucleic Acid Accession #: NM\_002851.1  
 Coding sequence: 148..7092

45 1 11 21 31 41 51  
 | | | | |  
 CACACATAGC CAGCAGCAT CTCACTTGA TCTATACACT GGAGGATTAA AACAAACAAA 60  
 CAAAAAATC ATTTCTCTCG CTCCTCTCC CTCTCCACTC TGAGAAGCAG AGGAGCCGCA 120  
 CGGCGAGGGG CGCAGACCG TCTGGAATG CGAATCTTAA AGCGTTTCTC CGCTTGCAAT 180  
 50 CAGCTCTCTC GTGTTTGC CGCTGATGG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240  
 CTGTTGAAG AGATTGGCTG GTCTATACA GGAGCACTGA ATCAAAAAA TTGGGGAAAG 300  
 AAATATCAA CATGTAATAG CCAAAACAA TCTCTATCA ATATTGATGA AGATCTTACA 360  
 CAAGTAATG TGAATCTTAA GAACTTAAA TTTCAGGGTT GGGATAAAC ATCATTGGAA 420  
 AACACATTCA TTCATAACAC TGGGAAACA GTGAAATTA ATCTCACTAA TGACTACCGT 480  
 55 GTCAGCGGAG GAGTTTCAGA AATGGTGTTT AAAGCAAGCA AGATAACTTT TCACTGGGGA 540  
 AAATGCAATA TGTACTCTGA TGATGAGAG CATAGTTTAG AAGGACAAA ATTTCCACTT 600  
 GAGATGCAA TCTACTGCTT TGATGCGGAC CGATTTTCAA GTTTTGAGGA AGCAGTCAAA 660  
 GGAAGAGGGA AGTTAAGAGC TTATCCATT TTGTTTGAAG TTGGGACAGA AGAAAAATTG 720  
 GATTTCAAG CGATTATTGA TGGAGTGAAG AGTGTAGTC GTTTTGGGAA CAGGCTGCT 780  
 60 TTAGATCCAT TCATACTGTT GAACCTTCTG CCAAACTCAA CTGACAAGTA TTACATTAC 840  
 AATGGCTCAT TGACATCTCC TCCTGACACA GACACAGTTG ACTGGATTGT TTTTAAAGAT 900  
 ACAGTTAGCA TCTCTGAAAG CCAAGTTGGT GTTTTGTG AAGTCTTAC AATGCAACAA 960  
 TCTGTATAG TCATGCTGAT GGAATACTTA CAAACAAAT TTGAGAGCA ACAGTACAAG 1020  
 65 TTCTCTAGAC AGGTGTTTTC CTCATACACT GGAAAGGAAG AGATTATGA AGCAGTTTGT 1080  
 AGTTCAGAAC CAGAAAATG TCAGGCTGAC CCAGAGAATT ATACCAGCCT TCTGTGTACA 1140  
 TGGGAAAGAC CTCGAGTCGT TTATGATACC ATGATTGAGA AGTTTGCAGT TTTGTACAG 1200  
 CAGTTGGATG GAGAGGACCA AACCAAGCAT GAATTTTGA CAGATGGCTA TCAAGACTTG 1260  
 GGTGCTATT TCAATAATTT GCTAOCATAT ATGAGTTATG TTCTTCAGAT AGTAGCCATA 1320  
 TGCACTAATG GCTTATATG AAAATACAGC GACCAACTGA TTGTGACAT GCCTACTGAT 1380  
 70 AATCTGAAAC TTGATCTTT CCTGGAATTA ATTGGAAGT AAGAAATAAT CAAGGAGGAG 1440  
 GAAGAGGAA AAGACATTGA AGAAGGCGCT ATTGTGAATC CTGTTAGAGA CAGTGTACA 1500  
 AACCAATCA GGAAGAAAGGA ACCCAGATT TCTACCACAA CACACTACAA TCGCATAGGG 1560  
 ACAGAAATCA ATGAAGCCAA GACTAACCGA TCCCAACAA GAGGAAGTGA ATTCTCTGGA 1620  
 75 AAGGGTATG TTCCCAATAC ATCTTTAAAT TCCACTTCCC AACCACTCAC TAAATTAGCC 1680  
 ACAGAAAGAT ATATTCTCT GACTTCTCAG ACTGTGACTG AACTGCCACC TCACACTGTG 1740  
 GAAGGTACTT CAGCCTCTTT AATGATGSC TCTAAACTCT TTCTTAGATC TCCACATATG 1800  
 AACTGTGCGG GGAAGTCAAG ATCCTTAAAT ACAGTTTCTA TAACAGAATA TGAGGAGGAG 1860  
 AGTTTATGTA CAGTTTCAA GCTTGATACT GGAGCTGAAG ATTCTTCAGG CTCAGTCCC 1920  
 GCAACTCTG CTATCCCAT CATCTCTGAG AACATATCCC AAGGGTATAT ATTTCTCTCC 1980  
 80 GAAACCCAG AGACAATAAT ATATGATGTC CTTATACAG AATCTGCTAG AAATGCTTCC 2040  
 GAAGATTCAA CTTCTACAG TFCAGAGAA TCACTAAAGG ATCCTTCTAT GGAGGAAAT 2100  
 GTGTGTTTC CTAGCTCTAC AGACATAACA GCACAGCCCG ATGTTGGATC AGGCAGAGAG 2160  
 AGCTTCTTCC AGACTAATTA CACTGAGATA CGTGTGATG AATCTGAGAA GACACCAAG 2220  
 TCCCTTCTG CAGGCCCAGT GATGTACAG GGTCCCTCAG TTACAGATCT GGAAATGCCA 2280  
 85 CATTATTCTA CTTTGCCTA CTTCCCACT GAGGTAAAC CTTCTGCTTT TACCCCATCC 2340  
 TCCAGACAA AGGATTGTGT CTCCAGGTC AACGTGGTAT ACTGCGAGAC AACCCACCG 2400  
 GTATACAATG GTGAGACACC TCTTCAACCT TCCTACAGTA GTGAAGTCTT TCCTTAGTCT 2460  
 ACCCTTTGTG TGCTTGACAA TCAGATCTCT AACACTAGCC CTGCTGCTTC AAGTAGTGAT 2520

	TOGGCCTTGC	ATGCTACGCC	TGTATTTCCC	AGTGTGATG	TGTCAATTGA	ATCCATCCTG	2580
	TCTTCTTATG	ATGGTGACCC	TTTGCTTCCA	TTTTCTCTG	CTTCTTCAG	TAGTGAATTG	2640
	TTTGCCCATC	TGCATACAGT	TTCTCAAATC	CTTCCACAAG	TTACTTCAGC	TACCGAGAGT	2700
5	GATAAGGTGC	CCTTGATCGC	TTCTCTGCCA	GTGGCTGGGG	GTGATTGTCT	ATTAGAGCCC	2760
	AGCCTTGCTC	AGTATTCTGA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GACGCTGGAA	2820
	TTTGTGATG	AATCTGGTGT	TCTTTATAAA	ACGCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTGATG	CCATGATGCA	TGCAGCTTCT	TCAGGGCCTG	AACCTTCTTA	TGCCTGTGCT	2940
	GATAATGAGG	GCTCCCAACA	CATCTTCACT	GTCTCTTACA	GTCTGCAAT	ACCTGTGSCAT	3000
10	GATTCTGTGG	GTGTAACTTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGTGGGG	AATGCTCTGG	AGCCTCTTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACTT	GTCTCTGTAG	CTGAATTTAC	ATATACAAACA	3240
	TCGTGTTTGG	GTGATGATAA	TAAAGGCGCT	TCTAAAAGTG	AAATAATATA	TGGAAATGAG	3300
15	ACTGAACATG	AAATTCCTTC	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCATG	3360
	CCCAACATGT	ATGATAATGT	AAATAAGTTG	AATGCGTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTCTTAGCA	CCAAGGGCAT	GTTCCTCAGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAAT	AACCTTTTCA	TTCAACCTAC	ACATACTGTC	3540
	TCTCAAGCAT	CTGGTGACAG	TTGCTTAAAT	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
20	TCCTCTGACC	CTGCTTCTAG	TGAAATGTTA	TCTCCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
	ACCTCAGCTT	CTTTTAGTAG	TGAAGTATTG	CTACAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAAACATG	TCTTCCAGCT	GTGCCCAAGT	ATCCAATATT	GCTTGAAACC	3780
	CCCAAGTTGG	ATAAAATTAG	TTCTACAATG	TTGATCTTCA	TTCTGCTTCA	TTCTGCTTCA	3840
	AGTGAACAAA	TGCTGACCTC	TACATCTGTA	CCAGTTTTTG	ATGTGTGCGC	TACTTCTCAT	3900
25	ATGCACTCTG	CTTCACTTCA	AGGTTTGACC	ATTTCTTATG	CAAGTGAGAA	ATATGAACCA	3960
	GTTTTGTAA	AAAGTGAAGG	TTCCCAACAA	GTGGTACCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAAC	CAGGCCCATC	CCCCAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAGCTTATA	4140
	CATTCCGATG	AAATTTTAAAC	CTCCACCAAA	AGTTCTGTTA	CTGGTAAGGT	ATTTGCTGGT	4200
30	ATTCCAACAG	TGTGTTCTGA	TACATTGTGA	TCTACTGATC	ATTCTGTTC	TATAGGAAAT	4260
	GGGCATGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTCTGTATAC	CTCAACAAAG	4320
	TTGCTGTTTC	CTTCTAAGGC	AACCTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCCGGT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGACAGA	4440
	GATAGTGATG	GCTTATCCAT	TCATAAGTGT	ATGTGATGCT	CTTCTATAG	AGAATCACAG	4500
35	GAAAAGGTAA	TGAATGATTG	AGACACCCAC	GAAAACAGTC	TTATGGATCA	GAATAATCCA	4560
	ATCTCATACT	CATCATCTGA	GAATTCGTAA	GAAGATAATA	GAGTCACAAG	TGATCTCTCA	4620
	GACAGTCAAA	CTGGTATGGA	CAGAAGTCCCT	GGTAAATCAC	CATCAGCAAA	TGGGCTATCC	4680
	CAAAAGCACA	ATGATGGAAG	AGAGGAAAT	GACATTCAGA	CTGGTAGTGC	TCTGCTTCTT	4740
	CTCAGCCCTG	AAATCTAAGC	ATGGGCAGTT	CTGACAAGTG	ATGAAGAAAG	TGGATCAGGG	4800
40	CAAGGTACCT	CAGATAGCCT	TAATGAGAAT	GAGACTTCCA	CAGATTTCAG	TTTTGCAGAC	4860
	ACTAATGAAA	AAAGTCTGTA	TGGGATCCTG	GCAGCAGGTG	ACTCAGAAAT	AACTCCTGGA	4920
	TTCCCAAGT	CCCCAACATC	ATCTGTTACT	AGCGAGAACT	CAGAAGTGTG	CCACGTTTCA	4980
	GAGGCAGAGG	CCAGTAATAG	TAGCCATGAG	TCTCGTATTG	GTCTAGCTGA	GGGGTTGGAA	5040
	TCCGAGAAGA	AGGCAGTTAT	ACCCCTTGTT	ATCGTGTGCT	CCCTGACTTT	TATCTGTCTA	5100
45	TGGGTTCTTG	TGGGTATTCT	CATCTACTGG	AGGAAATGCT	TCCAGACTGC	ACACTTTTAC	5160
	TTAGAGGACA	GTACATCCCC	TAGAGTTATA	TCCACACCTC	CACACCTTAT	CTTTCCAATT	5220
	TCAGATGATG	TCCGAGCATG	TCCAATAAAG	CACCTTTCCA	AGCATGTTGC	AGATTACAT	5280
	GCAAGTAGTG	GTTTACTAGA	AGAATTGAG	ACACTGAAAG	AGTTTACCA	GGAAGTCAG	5340
	AGCTGTACTG	TTGACTTAGG	TATTACAGCA	GACAGCTCCA	ACCACCCAGA	CAACAGCAC	5400
50	AAGAATCGAT	ACATAAATAT	CGTTGCTTAT	GATCATAGCA	GGTTAAGCT	AGCACAGCTT	5460
	GCTGAAAAGG	ATGGCAAACT	GACTGATTAT	ATCAATGCCA	ATTATGTTGA	TGGCTACAA	5520
	AGACCAAAAG	CTTATATTGC	TGCCCAAGGC	CCACTGAAAT	CCACAGCTGA	AGATTCTTGG	5580
	AGAATGATAT	GGGAACATAA	TGTGGAAGTT	ATTGTGATGA	TAACAAACCT	CGTGAGAGAA	5640
	GGAGGAGAA	AATGTGATCA	GTACTGGCCT	GCCGATGGGA	GTGAGGAGTA	CGGGAACCTT	5700
55	CTGGTCACTC	AGAAGAGTGT	GCAAGTGCTT	GCCTATTATA	CTGTGAGGAA	TTTTACTCTA	5760
	AGAAACACAA	AAATAAAAAA	GGGCTCCAG	AAAGGAAGAC	CCAGTGAGCG	TGTGCTCACA	5820
	CAGTATCACT	ACACGCACTG	GCTGACATG	GGAGTACCAG	AGTACTCCCT	GCCAGTGCTG	5880
	ACCTTTGTGA	GAAAGGCAGC	CTATGCAAG	CGCCATGCA	TGGGGCCTGT	TGTCTGTCAC	5940
	TGCACTGCTG	GAGTTGGAAG	AACAGGCACA	TATATTGTGC	TAGACAGTAT	GTGTCAGCAG	6000
60	ATTCAACAGC	AAGGAAGTGT	CAACATATTT	GGCTTCTTAA	AACACATCCG	TTCAACAAGA	6060
	AATTATTGTT	TACAACTGGA	GGAGCAATAT	GTCTTCACTT	ATGATACACT	GGTTGAGGCC	6120
	ATACTTAGTA	AAGAACTGGA	GGTGCTGGAC	AGTCATATTC	ATGCCATGTT	TAATGCACTC	6180
	CTCATCTCTG	GACCAGCAGG	CAAAACAAAG	CTAGAGAAAC	AATTCACGCT	CCTGAGCCAG	6240
	TCAAATATAC	AGCAGAGTGA	CTATTCTGCA	GCCCTAAAGC	AATGCAACAG	GGAAAGAAT	6300
65	CGAACTTCTT	CTATCATCCC	TGTGGAAAGA	TCAAGGGTTG	GCATTTCACT	CCTGAGTGGA	6360
	GAAGGCACAG	ACTACATCAA	TGCCCTCCTAT	ATCATGGGCT	ATTACAGAG	CAATGAATTC	6420
	ATCATTACCC	AGCACCTCTC	CCTTCATACC	ATCAAGGATT	TCTGGAGGAT	GATATGGGAC	6480
	CATAATGCC	AACTGGTGGT	TATGATTCCT	GATGGCCAAA	ACATGGCAGA	AGATGAATTT	6540
	GTTTACTGGC	CAAAATAAGA	TGAGCCTATA	AATTTGTGGA	GCTTTAAGGT	CACCTTTATG	6600
70	GCTGAAGAAC	ACAAATGCTT	ATCTAATGAG	GAAAACTTAA	TAATTTCAGGA	CTTTATCTTA	6660
	GAACTCACAC	AGGATGATTA	TGTACTTGAA	GTGAGGCAC	TTCAAGTGTCC	TAAATGGCCA	6720
	AATCCAGATA	GCCCCATTAG	TAAAACCTTT	GAACTTATAA	GTGTTATAAA	AGAAGAAGCT	6780
	GCCAATAGGG	ATGGGCTTAT	GATTGTTTAT	GATGAGCATG	GAGGAGTGAC	GGCAGGAAC	6840
	TTCTGTGCTC	TGACAACCTC	TATGCACCAA	CTAGAAAAG	AAAATTCCGT	GGATGTTTAC	6900
75	CAGGTAGCCA	AGATGATCAA	TCTGATGAGG	CCAGGAGTCT	TGTCTGACAT	TGAGCAGTAT	6960
	CAGTTTCTCT	ACAAAGTGTG	CCTCAGCCTT	GTGAGCACAA	GGCAGGAAGA	GAATCCATCC	7020
	ACCTCTCTGG	ACAGTAATGG	TGCAGCATTG	CCTGATGGA	ATATAGCTGA	GAGCTTAGAG	7080
	TCTTTAGTTT	AACACAGAAA	GGGGTGGGGG	GACTCACATC	TGAGCATTGT	TTTCTCTTTC	7140
	CTAAAATTAG	TCAGGAAAAAT	CAGTCTAGTT	CTGTTATCTG	TTGATTTCCT	ATCACCTGAC	7200
80	AGTAACTTTC	ATGACATAGG	ATTCTGCGCG	CAAAATTATA	TCATTAAACA	TGTGTGCTTT	7260
	TTTGCAAGAC	TGTAAATTTA	CTTATTATGT	TTGAACTTAA	ATGATTGAAT	TTTACAGTAT	7320
	TTCTAAGAA	GGAATTGTGG	TATTTTTTTC	TGTATTGATT	TTAACAGAAA	ATTTCAATTT	7380
	ATAGAGGTTA	GGAATTCCAA	ACTACAGAAA	ATGTTTGTGT	TTAGTGTCAA	ATTTTATAGT	7440
	GTATTGTGAG	CAATTATCAG	GTTTGTGAGA	AATATAACTT	TTAATACAGT	AGCCTGTAAA	7500
	TAAACACTCT	TTCCATATGA	TATTCAACAT	TTTACAACCT	CAGTATTTC	CTAAGTAGAA	7560
85	AATAATCTGT	TACTTATGTT	AAATACTGCC	CTAGTGTCTC	CATGGACCAA	ATTTATATTT	7620
	ATAATTGTAG	ATTTTATAT	TTTACTACTG	AGTCAAGTTT	TCTAGTTCTG	TGTAATTGTT	7680
	TAGTTTAATG	ACGTAGTTCA	TTAGCTGGTC	TTACTCTACC	AGTTTCTGGA	CATTGTATTG	7740

WO 02/086443

PCT/US02/12476

TGTTACCTAA GTCATTAACT TTGTTTCAGC ATGTAATTTT AACTTTTGTG GAAAAAGAA 7800  
 ATACCTTCAT TTGAAAGAA GTTTTTATGA GAATAACACC TTACCAACA TTGTTCAAAT 7860  
 GGTTTTTATC CAAGGAATTG CAAAAATAA TATAATATT GCCATTAAAA AAAAAAATAA 7920  
 AAAAAAATAA AAAAAAATAA A

Seq ID NO: 583 Protein sequence  
 Protein Accession #: NP\_002842.1

1 11 21 31 41 51  
 MRILKRLFLAC IQLLCVCRID WANGYYRQR KLVVEIGWSY TGALEQKNWG KKYPTCNSPK 60  
 QSPINIDEDL TQVNVNLKRL KFGQMDKTSI ENTFIHNTGK TVEINLTNDY RVSGGVSEMV 120  
 FKASKITFMW GKCNMSDDGS EHSLEGQKFP LEMQIYCFDA DRPSSFEZAV KKGKLRALS 180  
 ILFEVGTEN LDFKAIIDGV ESVSFRGKQA ALDPFILLNL LPNSTDKYYI YNGSLTSPPC 240  
 TDIVDVIVFK DTVISIESQL AVFCEVLTMQ QSGVYMLMDY LQNNFREQQY KFSRQVFSY 300  
 TGKEEIEHAV CSSEPEVQA DPENYTSLLV TWERPRVVDY TMIEKFAVLY QQLDGEDQTK 360  
 HEFLTDSYQD LGAILNLNLP NMSYVLQIVA ICTNGLYGKY SDQLIVDMPT DNEFLDLFPE 420  
 LIGTEIEIKE EEEGKDIEBG AIVNPGRDSA TNQIRKKEPQ ISTTTHYNRI GTRYNEAKTN 480  
 RSPTRGSEPS KGQDVNPTSL NSTSQPVTKL ATEKDILSTS QTVELPPHT VEGTSASLND 540  
 GSKTVLRSPH MNLSGTAESL NTVSIYEYEE ESLLTSFKLD TGAEDSSGSS PATSAIPFIS 600  
 ENISQGYIFS PKNVDKISST VLIPESARNA SEDSTSSGSE ESLKDPSEMG NVWFPSSTDI 660  
 TAQPDVGGGR ESFLQNTYIE IRVDESEKTT KSFAGPVMS QGPSVTDLEM PHYSTFAYFP 720  
 TEVTPHAFPT SSRQQLVST VNVVYSQTTQ PVYNGETPLQ PSYSSEVFPL VTFLLLDNQI 780  
 LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAFLI PFSSASPSSE LFRHLHTVSQ 840  
 ILPQVTSATE SDKVLPHASL PVAGDGLLE PSLAQYSDVL STTHAASETL EPGSESGVLY 900  
 KTLMFQVSEP PSSDAMMHAR SSGPEPSYAL SDNEGSQHIF TVSYSSAIPV HDSVGVITYG 960  
 SLFSGSHIP IPKSSLITPT ASLLQPTHAL SGDGENSGAS SDSEFLLPDT DGLTALNISS 1020  
 PVSVAEFTYT TSVFGDDNKA LSKSEIIVGN ETELQIPSEF EMVYSESTV MPNMYDNVKN 1080  
 LNASLQETSV SISSTKGMFP GSAHATTKV FDREISQVPE NNFVSQPTHT VSQASGDTSL 1140  
 KPVLSANSEP ASSDPASSE LSPSTQLLFY ETSASFSTEV LLQPSFQASD VDTLLKTVLP 1200  
 AVPSDILVLE PKNVDKISST MLHLIVNSA SSENMLESTS VPVFDVSPIS HMHSASLQGL 1260  
 TISYASEKYE PVLLKSESSH QVPSLSYEND ELFTQANLEI NQAEPPKGRH VFATPVLSID 1320  
 EPLNTLNLKL IHSDEILTST KSSVTGKVA GIPTVASDTF VSTDHSPVIG NGHVAITAVS 1380  
 PHRDGSVTST KLLFPSSKATS ELSSSAKSDA GLVGGGEGD DDDGDDDDDD RDSGLSLHK 1440  
 CMSCSYRES QKVMNDSDT HENSLMDQNN PISYBLSNS EEDNRVTSVS SDSQTGMDRS 1500  
 PGKSPSANGL SQKNDGKKEE NDQTSALL PLSPSKAWA VLTSDEESGS GQGTSDSLNE 1560  
 NETSDPFSFA DTNEKDADGI LAAGDSEITP GFFQPSSTSS TSENSEVFHV SEAEASNSSH 1620  
 BSRIGLAEBL ESSKKAIVPL VIVSALTPIE LVLVLGILIY WRKCFQTAHF YLEDSTSPRV 1680  
 ISTPPTPIFP ISDDVGAIP KHPFKHVDL HASSGTEEF ETLKEFYQEV QSCCTVDLGIT 1740  
 ADSNHPDNK HKRYINIVA YHSRVLQAL LAEKDGKLD YINANYVDGY NRPKAYIAAQ 1800  
 GPLKSTAEFD WRIWEHNV VIVMITNLVE KGRKCDQYV PADGSEEGYN FLVTKQSVQV 1860  
 LAYTYVRNFT LRNTKIKKGS QKGRPSGRVV TQYHYTQWPD MGVPYSLPV LTFVRAKAYA 1920  
 KRHAGFPVVV HCSAGVGRFG TYIVLDSMLQ QIQHEGTVNI FGFLKHRSQ RNYLVQTEEG 1980  
 YVFIHDTLVE AILSKETBVL DSHIHAYVNA LLIPGPAGTK KLEKQFOLLS QSNIQSDYS 2040  
 AALKQCNREK NRTSSIIPE RSRVGISSLS GEGTDYINAS YIMGYYSNE FIITQEPHLLH 2100  
 TIKDFWRMIV DINAQLVMI PDGQNMARDE FVYWPNDPE INCBSFKVTL MAEHEKCLSN 2160  
 EEKLIQDFI LEATQDDVYL EVRFQCPKW PNPDSPIKST FELISVIEKE AANRDGPMIV 2220  
 HDEHGGVTAG TFCALTLMH QLEKENSVDV YQVAKMINLM RGVVFADIEQ YQFLYKVLIS 2280  
 LVSTRQENP STSLDSNGAA LPDGNIAESL ESLV

Seq ID NO: 584 DNA sequence  
 Nucleic Acid Accession #: NM\_005688.1  
 Coding sequence: 126..4439

1 11 21 31 41 51  
 CCGGGCAGST GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGCGCGCGTT GTCTGGAGC 60  
 AGGGGCGCAG GAATCTCTGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCTGAGC 120  
 AGAAGATGAA GGATATCGAC ATAGGAAAG AGTATATCAT CCCAGTCCT GGTATAGAA 180  
 GTGTGAGGGA GAGAACCCAG ACTTCTGGGA CGCACAGAGA CGTGAAGAT TCCAAGTTCA 240  
 GGAGAACTCG ACCGTTGGAA TGCCAGATG CTTTGGAAAC AGCAGCCCGA GCCAGGGGCC 300  
 TCTCTCTTGA TGCTCTCAATG CATCTCAGC TCAGAACTCT GGATGAGGAG CATCCCAAGG 360  
 GAAAGTACCA TCATGGCTTGA AGTGCTCTGA AGCCCATCCG GACTACTTCC AAACACCCAG 420  
 ACCAGTGGGA CAATGCTGGG CTTTCTTCTT GTATGACTTT TCTGTTGGCT TCTTCTCTGG 480  
 CCGGTGTGGC CCACAGAGAG GGGGAGCTCT CAATGGAAGA CGTGTGGTCT CTGTCCAAGC 540  
 ACGAGTCTTC TGACGTGAAC TGCAGAAGAC TAGAGAGACT GTGCAAGAA GAGCTGAATG 600  
 AAGTGTGGCC AGACGCTGCT TCCCTGCGAA GGGTTGTGTG GATCTTCTGC CGCACCAGGC 660  
 TCATCTCTGC CATCGTGTGC CTGATGATCA CGCAGCTGGC TGGCTTCAGT GGACCAGCCT 720  
 TCATGTGTGA ACACCTCTTG GAGTATACCC AGGCAACAGA GTCTAACCTG CAGTACAGCT 780  
 TGTGTGTAGT GCTGGGCTCT CTCTGACGG AAATCGTGGC GTCTTGTGCG CTTGCACTGA 840  
 CTGTGGCATT GAATTACCGA ACCGGTGTCC GCTTGGCGGG GGCCATCTTA ACCATGGCAT 900  
 TTAAGAAAGT CCTTAAGTTA AAGAAACATTA AAGAGAAATC CTTGGGTGAG CTCATCAACA 960  
 TTTGTCTCAA CGATGGGCGAG AGAATGTTTG AGGCAGCAGC CGTGGGCGAG CTGCTGGCTG 1020  
 GAGGACCCGT TGTGTCCATC TTAGGCATGA TTTATAATGT AATTATTCTG GGACCAACAG 1080  
 GCTTCTCTGG ATCAGCTGTT TTTATCTCTT TTTACCTAGC AATGATGTTT GCATCAACGGC 1140  
 TCACAGCATA TTTACAGAGA AAATGCTGTG CCGCCACGGA TGAACGTGTC CAGAAGATGA 1200  
 ATGAAGTTCT TACTTACATT AAATTTATCA AAATGTATGC CTGGGTCAA GCAATTTCTC 1260  
 AGAGTGTTC AAAAAATCCG GAGGAGGAGC GTCGATATT GGAAGAGGCC GGGTACTTCC 1320  
 AGGGTATCAC TGTGGGTGTG GCTCCCATTT TGTGTGTGAT TGCCAGCGTG GTGACCTTCT 1380  
 CTGTCTATAT GACCTTGGG TTAGATCTGA CAGCAGCACA CGCTTTTACA GTGGTGACAG 1440  
 TCTTCAATTC CATGACTTTT GCTTTGAAG TAACACCGTT TTCAGTAAAG TCCCTCTCAG 1500  
 AAGCCTCAGT GCTGTGTGAC AGATTTAAGA GTTGTGTTCT AATGGAAGAG GTTCACATGA 1560  
 TAAGAACCAA ACCAGCCAGT CCTCACATCA AGATAGAGAT GAAAAATGCC ACCTTGGCAT 1620  
 GGGACTCTCT CCACTCCAGT ATCCAGAACT CGCCCAAGCT GACCCCAAAA ATGAAAAAAG 1680  
 ACAAGAGGCG TCCAGCGGAG AAGAAAGAGA AGGTGAGGCA GCTGCAGCGC ACTGAGCATC 1740  
 AGGCGGTGCT GGCAGAGCAG AAGGCGCACC TCCTCTCTGA CAGTGACGAG CGGCCAGTTC 1800  
 CCGAAGAGGA AGAAGCGAAG CACATCCACC TGGGCCACCT GCGCTACAG AGGACACTGC 1860

ACAGCATCGA TCTGGAGATC CAAGAGGGTA AACTGGTGG AATCTGCGGC AGTGTGGGA 1920  
 GTGGAAAAAC CTCTCTCATT TCAGCCATT TAGGCCAGAT GACGCTCTTA GAGGGCAGCA 1980  
 TGGCAATCAG TGGAACTTTC GCTTATGTGG CCCAGCAGGC CTGGATCCTC AATGCTACTC 2040  
 TGAGAGACA CATCTCTGTT GGAAGGAAT ATGATGAAGA AAGATACAAC TCTGTGCTGA 2100  
 5 ACAGCTGCTG CCGTAGGGCT GACCTGGCCA TTCTTCCCAG CAGCGACCTG ACGGAGATTG 2160  
 GAGAGCGAGG AGCCAACTCT AGCGGTGGGC AGCGCCAGAG GATCAGCCTT GCCCGGGCCT 2220  
 TGTATAGTGA CAGGAGCATC TACATCCTGG ACGACCCCTT CAGTGCCTTA GATGCCCATG 2280  
 TGGGCAACCA CATCTTCAAT AGTGCTATCC GGAACATCT CAAGTCCAAG ACAGTCTCTG 2340  
 10 TTGTTACCCA CAGTTACAGA TACCTGGTTG ACTGTGATGA AGTGTCTTC ATGAAAGAGG 2400  
 GCTGTATTAC GGAAGAGGC ACCCATGAGG AACTGATGAA TTTAAATGGT GACTATGCTA 2460  
 CCATTTTAA TAACCTGTGT CTGGGAGAGA CACCGCCAGT TGAGATCAAT TCAAAAAAGG 2520  
 AAACAGTGG TTCACAGAAG AAGTCACAAG ACAAGGGTCC TAAAAAGGA TCAGTAAAGA 2580  
 AGGAAAAAGC AGTAAAGCCA GAGGAAGGGC AGCTTGTGCA GCTGGAAGAG AAAGGGCAGG 2640  
 15 GTTCAGTCCC CTGGTCAGTA TATGGTGTCT ACATCCAGGC TGCTGGGGGC CCCTTGGCAT 2700  
 TCCTGGTTAT TATGGCCCTT TTCTATGCTGA ATGAGGCAG CACCGCCTTC AGCAGCTGGT 2760  
 GGTGAGTTA CTGGATCAAG CAAGGAAGCG GGAACACCA CTGACTATGCC AGCATCTACG 2820  
 CCTCGGTGAG TGACAGCATG AAGGCAATC CTATATGCA GTACTATGCC AGCATCTACG 2880  
 CCCTCTCCAT GGCAGTCTAT CTGATCCTGA AAGCCATCG AGGATGTGTC TTTGTCAAGG 2940  
 20 GCACGCTCGC AGCTTCTTCC CGGCTGATG ACGAGCTTTT CCGAAGGATC CTTGGAAGCC 3000  
 CTATGAAGTT TTTTACACG ACCCCACAG GAGGATTTCT CAACAGGTTT TCCAAAGACA 3060  
 TGGATGAAGT TGACGTGGCG CTGCCGTTCC AGGCCGAGAT GTTCATCCAG AACGTTATCC 3120  
 TGGTGTCTTT CTGTGTGGGA ATGATGCGAG GAGTCTTCCC GTGTTCTCTT GTGGCAGTGG 3180  
 GGGCCCTTGT CATCTCTTTT TCAGTCTGCG ACATGTCTCT CAGGTTCTCT ATTCCGGAGC 3240  
 25 TGAAGGCTCT GAGCAATATC ACGCAGTCAC CTTTCTCTC CCACATCACG TCCAGCATAC 3300  
 AGGGCCTTGC CACCATCCAC GCCTACAATA AAGGGCAGGA GTTCTGTCAC AGATACCAGG 3360  
 AGCTGCTGGA TGACAACAA GCTCCTTTT TTTGTGTTAC GTGTGCGATG CGGTGGCTGG 3420  
 CTGTGGGCTT GGACCTCATC AGCATCGCCC TCATCACCAC CACGGGCTG ATGATCGTTC 3480  
 TTAATGACGG GCAGATTTCC CCAGCCTATG CGGCTCTGCG CATCTCTTAT GCTGTCCAGT 3540  
 30 TAACGGGGCT GTTCCAGTTT ACGGTGAGAC TGGCATCTGA GACAGAAGCT CGATTACCTC 3600  
 CGGTGGAGAG GATCAATCAC TACATTAAGA CTCTGTCTCT GGAAGCACTT GCCAGAATTA 3660  
 AGAACAGGC TCCCTCCCTT GACTGGCCCC AGGAGGAGA GGTGACCTTT GAGAACGCG 3720  
 AGATGAGGTA CCGAGAAAA CTCCCTCTTG TCCTAAGAA AGTATCCTTC ACGATCAAC 3780  
 CTAAGAGAGA GATTGGCATT GTGGGGCGGA CAGGATCAGG GAAGTCTCTG CTGGGGATGG 3840  
 35 CCCTCTTCCG TCTGGTGGAG TTATCTGGAG GCTGCATCAA GATTGATGGA GTGAGATCA 3900  
 GTGATATGG CCTTGGCAGC CTCCGAAGCA AACTCTCTAT CATCTCTCAA GAGCCGGTGC 3960  
 TGTTCAGTGG CACTGTGACA TCAAATTTGG ACCCTTCAA CCACTACACT GAAGACCA 4020  
 TTTGGGATGC CCTGGAGAGG ACACACATGA AAGAATGTAT TGCTCAGCTA CCTCTGAAAC 4080  
 TTGAATCTGA AGTGATGGAG AATGGGGATA ACTTCTCAGT GGGGGAACGG CAGCTCTTGT 4140  
 40 GCATAGCTAG AGCCCTGCTC CGCCACTGTA AGATTCTGAT TTTAGATGAA GCCACAGCTG 4200  
 CCATGGACAC AGAGACAGAC TTATGTATTC AAGAGACCAT CCGAGAAGCA TTTGCAGACT 4260  
 GTACCATGCT GACCATTGCC CATCGCCTGC ACACGGTCTT AGGCTCCGAT AGGATTATGG 4320  
 TGCTGGCCCA GGGACAGGTG GTGGAGTTTG ACACCCCATC GGTCTCTCTG TCCAACGACA 4380  
 TGTCCGATTT CTATGCCATG TTTGCTGCTG CAGAGAACAA GGTGCTGTG AACGGCTGAC 4440  
 45 TCCTCCCTGT TGACGAAGTC TCTTTCTTT AGAGCATTCG CATTCCTGTC CTGGGGCGGG 4500  
 CCCTCATCG CGTCTCTCTA CCGAAACCTT GCCTTTCTCG ATTTTATCTT TCGCACAGCA 4560  
 TGTCCGGATT GCTCTGTGTT TTTCACTTTT AGGAGAGAGT ATATTTTGAT TATTGTATT 4620  
 ATTCCATATT CATGTAAACA AAATTTAGTT TTTGTTCTTA ATTGCACTCT AAAAGGTTCA 4680  
 50 GGGAAACGGT ATTATAATTG TATCAGAGGC CTATAATGAA GCTTTATACG TGTAGCTATA 4740  
 TCTATATATA ATTCTGTACA TAGCCTATAT TTACAGTGAA AATGTAAAGT GTTTATTTTA 4800  
 TATTAAAAAT AGCACTGTGC TAATAACAGT GCATATCTCT TTCTATCATT TTTGTACAGT 4860  
 TTGCTGTACT AGAGTCTGG TTTTGCTATT AGACTGTAGG AAGAGTAGCA TTTCTTCTT 4920  
 CTCTAGCTGG TGGTTTACG GTGCCAGGTT TTCTGGGTGT CCAAAGGAAG ACGTGTGCA 4980  
 ATAGTGGGCC CTCCGACAGC CCCCTCTGCC GCCTCCCCAC AGCGCTCCA GGGGTGGCTG 5040  
 55 GAGACGGGTG GCGGGCTGGA GACCATGCG AGCGCCGTGA GTTCTCAGG CTCTGCTCTT 5100  
 CTGTCTCTGT GTCACCTACT GTTCTGTGTA GGAGAGCAGC GGGGCGAAGC CCAGGCCCTT 5160  
 TTTCACTCCC TGCATCAAGA ATGGGGATCA CAGAGACATT CCTCCGAGCC GGGGAGTTTC 5220  
 TTTCTGCTCT TCTCTTTTTT GCTGTGTTT CTAACAAGA ATCAGTCTAT CCACAGAGAG 5280  
 TCCACTGCC CTAGGTTCTT ATGGCTGGCC ACTGCACAGA GCTCTCCAG TCCAAGACCT 5340  
 60 GTTGGTTCCA AGCCCTGGAG CCAACTGCTG CTTTGTGAGG TGGCACTTTT TCATTGCTCT 5400  
 ATTCCACAC CTCCACAGTT CAGTGGCAGG GCTCAGGATT TCGTGGGTCT GTTTTCTCTT 5460  
 CTCAACGAG TCGTCCGACA GTCTCTCTCT CTCTCTCCCC TCAAAGTCTG CAACCTTAAG 5520  
 CAGCTCTGC TAATCAGTGT CTCACTCTGG CGTAGAAGTT TTTGTACTGT AAAGAGACCT 5580  
 ACCTCAGGTT GCTGGTGTCT GTGTGGTTTG GTGTGTTCCC GCAAAACCCC TTTGTGCTGT 5640  
 65 GGGGCTGSTA GCTCAGGTGG GCGTGGTAC TGCTGTCTAT AGTTGAATGG TCAGGTTGCT 5700  
 ATGTCTGTAC CACTAGACA TTTCTGCTGC TTAGCATGTT TGCTGAACAC CTTGTGGAAG 5760  
 CAAAATCTG AAAATGTGAA TAAATTTATT TTGGATTG TAAAAAATA AAAAAAATA 5820  
 AAAAAAATA AAAAAAATA

Seq ID NO: 585 Protein sequence  
 Protein Accession #: NP\_005679.1

1 11 21 31 41 51  
 MKDIDIGKEY IIPSPGYSV RERTSTSGTH RDREDSKFRF TRPLEQDQAL ETAARAEGLS 60  
 75 LDASMSQLR ILDEHPKKG YHGLSALKP IRTTSKEQHP VDNAGLFSCM TFSWLSLAR 120  
 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQELNEV GPDAASLRV VWIFCRTRLI 180  
 LSVICLMTQ LAGPSGPAPM VKHLLTQA TESNLQYSL LVLGLLLEI VRSWSLALTW 240  
 ALNVRTGVL RLKALTMAFK KILKLNKE KSLGELINIC SNGQRMPEA AAVGSLLAG 300  
 80 FVVALIGLIY NVILLGPTGF LGSVAVFIFY PAMMFASRLT AYFRRCVAA TDERVQXNNE 360  
 VLTYYIKFIK YAWKAFPSQS VQKIREEERR ILEKAGYFQG ITVGVAPIV VIASVVTFSV 420  
 HMTLGFDLTA AQAFVTVTVF NSMTFALKVT PFSVKSLSA SVAVDKPKSL FLMEEVHMIK 480  
 NKPASPHIKI EMKNATLAW SSSHSIQNSP KLPKMKDK RASRGKKEKV RQLQRTHEQA 540  
 VLAQKGHLH LQSDERPSP EEEGKHHLG HRLQRTLHS IDLEIQEGL VIGCGSVGSG 600  
 85 KTSLSAILG QMTLLBGSIA ISGTFAYVAQ QAWILNATL DNILFGKEYD EERYNSVLNS 660  
 CCLRPDLAL PSSDLTEIGE RGNLSGGQR QRISLARALY SDRSIYILD PSALDAHV 720  
 NHIPNSAIRK HLKSKTVLFV THQLQYLVD DEVIFMKEGC ITERGTHEEL MNLNGDYATI 780  
 FNNLLGETP PVEINSKRET SGSQKKSQDK GPKTGSVKKE KAVKPEBQQL VQLEBKQGS 840

VPWSVYGVYI QAAGGGLAFL VIMALFMLNV GSTAFSTWVL SYWIKQSGSN TTVTRGNETS 900  
 VDSMKDNPH MQYYASIALY SMAVMLILKA IRGVVFPVKGT LRASSRLHDE LFRILRSPM 960  
 KFPDTPPTGR ILNRPFSKMD EVDVRLPFQA EMFIQNVILV FFCVGMIAGV FWFVLVAVGP 1020  
 LVILFVLIHI VSRVLIRELK RLDNITQSPF LSHITSSIQG LATIHAYNKG QEFLEHYQEL 1080  
 LDNDQAPFPL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HGQIPPAYAG LAISYAVQLT 1140  
 GLFQFTVRLA SETEARFTSV ERINHVIKTL SLEAPARIKN KAPSPDWQOE GEVTFENAEM 1200  
 RYRENLPVLV KKVSPFTIKPK EKIGIVGRTO SKSSSLGMAL FRLVELSGGC IKIDGVRISD 1260  
 IGLADLRSLK SIIPQEPVLF SGTVRSNLDP FNQYTEDQIW DALERTHEKE CIAQLPLKLE 1320  
 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETDLIQE TIREAFADCT 1380  
 MLTIAERLET VLGSDRIMVL AQGVVVEFDT PSVLLSNDSS RPYAMFAAAE NKVAVKGO

Seq ID NO: 586 DNA sequence  
 Nucleic Acid Accession #: NM\_001327.1  
 Coding sequence: 89..631

1 11 21 31 41 51  
 | | | | |  
 AGCAGGGGGC GCTGTGTGTA CCGAGAATAC GAGAATACCT CGTGGGCCCT GACCTTCTCT 60  
 CTGAGAGCCG GGCAGAGGCT CCGAGGCCAT GCAGGCCGAA GGCAGGGGCA CAGGGGGTTC 120  
 TAGGGGCCAT GCTGATGGCC CAGGAGGCCG TGGCATTCTT GATGGCCAG GGGGCAATGC 180  
 TGGGGGCCCA GAGAGGGCGG GTGCCACGGG CGGCAGAGGT CCGCGGGGCG CAGGGGCAGC 240  
 AAGGGCCCTG GGGCGGGGAG GAGGCGCCCC CGGGGTCCG CATGGCGGCG CGGCTTCAGG 300  
 GCTGAATGGA TGCTGCAGAT GCGGGGCCAG GGGGCCGAG AGCCGCTGCT TTGAGTTCTA 360  
 CCTCGCCATG CCTTTCGCGA CACCATGGA AGCAGAGCTG GCCCGCAGGA GCCTGGCCCA 420  
 GGATGCCCCA CCGCTTCCCG TGCCAGGGGT GCTTCTGAAG GAGTTCACTG TGTCCGGCAA 480  
 CATACTGACT ATCCGACTCA CTGCTGCAGA CCACCGCCAA CTGCAGCTCT CCATCAGCTC 540  
 CTGTCTCCAG CAGCTTTCCT TGTTGATGTG GATCAGCCAG TGCTTCTGCG CCGTGTCTTT 600  
 GGCTCAGCCT CCCTCAGGGC AGAGGCGCTA AGCCAGCCT GGGCGCCCTT CTAAGTTCAT 660  
 GCCTCTCTCC CTAGGGAATG GTCCAGCAC GAGTGGCCAG TTCATTGTGG GGGCCTGATT 720  
 GTTGTGCTCT GGAGGAGGAC GGCTTACATG TTGTGTTCTG TAGAAATAA AACTGAGCTA

Seq ID NO: 587 Protein sequence  
 Protein Accession #: NP\_001318.1

1 11 21 31 41 51  
 | | | | |  
 MQAEGSGTGG STGDADGPGG PGIPDGPGEN AGGPGEAGAT GGRGPRGAGA ARASGPGGGA 60  
 PRGPHGGAAS GLNGCCRCGA RGPESRLLEF YLAMPFATPM EAELARRSLA QDAPPLFPVG 120  
 VLLKEFTVSG NILTIRLTAA DHRQLQLSIS SCLQQLSLLM WITQCPLPVF LAQPPSGQRR

Seq ID NO: 588 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 52..459

1 11 21 31 41 51  
 | | | | |  
 CCTGTGGGCG CTTGACCTTC TCTCTGAGAG CCGGGCAGAG GCTCCGAGGC CATGCAGGCC 60  
 GAAGGCCAGG GCACAGGGCG TCGACGGGCG GATGCTGATG GCCCAGGAGG CCTGGCATT 120  
 CCTGATGGCC CAGGGGGCAA TGCTGGGGCG CCAGGAGAGG CGGGTGCCAC GGGCGGCAGA 180  
 GGTCCCGGGG GCGCAGGGCG AGCAAGGGCC TCGGGGCCGA GAGGAGGGCG CCGCGGGGT 240  
 CGCATGCGCG GTCCCGCTTC TGCGCAGGAT GGAAGGTGCC CCTCGGGGCG CAGGAGGCCG 300  
 GACAGCCGCC TGCTTCACTT CCGACTGACT GCTGCAGACC ACCGCCAACT GCAGCTCTCC 360  
 ATCAGCTCTT GTCTCCAGCA GCTTCCCTCG TTGATGTGGA TCACGCAGTG CTTTCTGCC 420  
 GTGTTTTTGG CTCAGGCTCC CTCAGGGCAG AGGCGCTAAG CCGAGCCTGG CGCCCTCTCC 480  
 TAGGTCATGC CTCCTCCCTT AGGGAATGGT CCGAGCAGCA GTGCCAGTT CATTGTGGGG 540  
 GCCTGATTGT TTGTGCTGG AGGAGGACGG CTTACATGTT TGTTCCTGTA GAAATAAAG 600  
 CTGAGCTA

Seq ID NO: 589 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 | | | | |  
 MQAEGSGTGG STGDADGPGG PGIPDGPGEN AGGPGEAGAT GGRGPRGAGA ARASGPRGGA 60  
 PRGPHGGAAS AQDGRCPGCA RRPDSRLQLF RLTAADHRQL QLSISSCLQQ LSLLMWITQC 120  
 FLVFLAQAP SGQRR

Seq ID NO: 590 DNA sequence  
 Nucleic Acid Accession #: NM\_005562.1  
 Coding sequence: 90..3671

1 11 21 31 41 51  
 | | | | |  
 ACAGCGGAGC GCAGAGTGAG AACCAACCAAC CGAGGCGCGG GGCAGCGACC CTGCAGCGG 60  
 AGACAGAGAC TGAGCGGCCG GCACCGGCCA TGCTGCGCT CTGGCTGGGC TGCTGCCTCT 120  
 GCTTCTCGCT CCTCTGCCG GCAGCCGGGG CCACTCCAG GAGGGAAGTC TGTGATTGCA 180  
 ATGGGAAGTC CAGGCAGTGT ATCTTTGATC GGGAACTTCA CAGACAACT GGTAAATGGAT 240  
 TCGCTGCTCT CAACTGCAAT GACAACTGT ATGGCATTCA CTGCAGAGAG TGCAAGAATG 300  
 GCTTTTACCG GCACAGAGAA AGGGACCGCT GTTTCGCCCT CAATTGTAACT TCCAAGGTT 360  
 CTCTTAGTGC TCGATGTGAC AACTCTGGAC GGTGCAGCTG TAAACCAAGT GTGACAGGAG 420  
 CCGATGCGGA CCGATGTCTG CCAGGCTTCC ACATGCTCAC GGATGCGGGG TGACCCCAAG 480  
 ACCAGAGACT GCTAGACTCC AAGTGTGACT GTGACCCAGC TGGCATGCA GGGCCCTGTG 540  
 ACGCGGGCCG CTGTGTCTGC AAGCCAGCTG TTAGTGGAGA ACGCTGTGAT AGGTGTGAT 600  
 CAGGTACTTA TAATCTGGAT GGGGGGAACC CTGAGGCGTG TACCCAGTGT TTCTGCTATG 660  
 GGCATTACAG CAGCTGCCGC AGCTCTGCAG AATACAGTGT CCATAAGATC ACCTCTACCT 720  
 TTCATCAGA TGTGATGGC TGGAGGCTG TCCAACGAAA TGGGTCTCCT GCAAAGCTCC 780  
 AATGTCACA GCGCCATCAA GATGTGTTTA GCTCAGCCCA ACGACTAGAC CCGTCTATT 840

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70  
 75

TTGTGGCTCC TGCCAAATTT CTGGGAATC AACAGGTGAG CTATGGGCAA AGCCTGTCCT 900  
 TTGACTACCG TGTGGACAGA GGAGGCAGAC ACCCATCTGC CCATGATGTG ATTCTGGAAG 960  
 GTGCTGGTCT ACGGATCACA GCTCCCTTGA TGCCACTTGG CAAGACACTG CCTTGTGGGC 1020  
 TCACCAGAGC TTACACATTC AGGTTAAATG AGCATCCAAG CAATAATTGG AGCCCCAGC 1080  
 TGAGTTACTT TGAGTATCGA AGGTTACTGC GGAATCTCAC AGCCCTCCGC ATCCGAGCTA 1140  
 CATATGGAGA ATACAGTACT GGGTACATTC ACAATGTGAC CCTGATTTCA GCGCCCTTGG 1200  
 TCTCTGGAGC CCCAGCACCC TGGGTGAAC AGTGTATATG TCCTGTGGGG TACAAGGGGC 1260  
 AATTCTGCCA GGATTTGTCT TCTGGCTACA AGAGAGATTC AGCGAGACTG GGGCCTTTTG 1320  
 GCACCTGTAT TCCTTGTAACT TGTCAAGGGG GAGGGGCTGT TGATCCAGAC ACAGGAGATT 1380  
 GTTATTTCAGG GGATGAGAAT CCTGACATTG AGTGTGCTGA CTGCCCAAT GTTCTTCTACA 1440  
 ACGATCCGCA CGACCCCGCG AGCTGCAAGC CATGTCCCTG TCATAACCGG TTCAGCTGCT 1500  
 CAGTGTGCCC GGAGACGGAG GAGGTGGTGT GCAATAACTG CCTCCCGGGG GTCACCGGTG 1560  
 CCGCTGTGA GCTCTGTCT GATGGCTACT TTGGGGACCC CTTTGGTGAA CATGGCCGAG 1620  
 TGAGGCTTGG TCAGCTGTGT CAATGCAACA ACAATGTGGA CCCCAGTGCC TCTGGGAATT 1680  
 GTGACCGGCT GACAGGCAGG TGTTTGAAGT GTATCCACAA CACAGCCGGC ATCTACTCGG 1740  
 ACCAGTGCAG AGCAGGCTAC TTGGGGGACC CATTGGCTCC CAACCCAGCA GACAAGTGTC 1800  
 GAGCTTGCAG CTGTAACCCC ATGGGCTCAG AGCCTGTAGG ATGTGCAAGT GATGGCACCT 1860  
 GTGTTTGCAG GCCAGGATT GTTGGCCCA ACTGTGAGCA TGGAGCATT AGCTGTCCAG 1920  
 CTGTGTATAA TCAGTGAAG ATTGAGATGG ATCAGTTTAT GCAGCAGCTT CAGAGAAATG 1980  
 AGGCCCTGAT TTCAAAGGCT CAGGCTGGTG ATGAGTAGT ACCTGATACA GAGCTGGAAG 2040  
 GCAGGATGCA GCAGGCTGAG CAGGCCCTTC AGGACATTCT GAGAGATGCC CAGATTTCAG 2100  
 AAGGTGCTAG CAGATCCCTT GGTCTCCAGT TGGCCAAAGT GAGGAGCCAA GAGAACAGCT 2160  
 ACCAGAGCCG CCTGGATGAC CTCAGATGA CTGTGGAAAG AGTTCGGGCT CTGGGAAGTC 2220  
 AGTACCAGAA CCGAGTTGAG GATACTCACA GGCTCATCAC TCAGATGCAG CTGAGCCTGG 2280  
 CAGAAAGTGA AGCTTCTCTG GGAACACATA ACATTCCTGC CTCAGACCCAC TACGTGGGGC 2340  
 CAAATGGCTT TAAAGTCTG GCTCAGGAGG CCACAAGATT AGCAGAAAGC CACGTTGAGT 2400  
 CAGCCAGTAA CATGGAGCAA CTGACAAGGG AAACAGGAGA CTATTCCAAA CAAGCCCTCT 2460  
 CACTGTGCGC CAAGGCCCTG CATGAAGGAG TCGGAAGCGG AAGCGTAGC CCGGACGGTG 2520  
 CTGTGTGCGA AGGGCTTGTG GAAAAATTGG AGAAAAACAA GTCCCTGGCC CAGCAGTTGA 2580  
 CAAGGGAGGC CACTCAAGCG GAAATTGAAG CAGATAGGTC TTATCAGCAC AGTCTCCGCC 2640  
 TCCTGGATTG AGTGTCTCGG CTTCAGGAG TCAGTGATCA GTCTTTTCTG GTGGAAGAAG 2700  
 CAAAGAGGAT CAACAAAAA GCGGATTAC TCTCAACGCT GGTAAACAGG CATATGGATG 2760  
 AGTTCAAGCG TACACAAAAA AATCTGGGAA ACTGGAAAGA AGAAGCACAG CAGCTCTTAC 2820  
 AGAATGGAAA AAGTGGGAGA GAGAAATCAG ATCAGCTGCT TTCCGTGCCC AATCTTGCTA 2880  
 AAAGCAGAGC ACAAGAGACA CTGAGTAGG GCAATGCCAT TTTTATGAA GTTGAGAGCA 2940  
 TCCTTAAAAA CCTCAGAGG TTTGACCTGC AGGTGGACAA CAGAAAAAGC GAAGCTGAAG 3000  
 AAGCCATGAA GAGACTCTCC TACATCAGCC AGAAGGTTTC AGATGCCAGT GACAAGACCC 3060  
 AGCAAGCAGA AAGAGCCCTG GGGAGGCGTG CTGCTGATGC ACAGAGGGCA AAGAATGGGG 3120  
 CCGGGGAGGC CCTGGAATC TCCAGTGAGA TTGAACAGGA GATTGGGAGT CTGAACCTGG 3180  
 AAGCCAATGT GACAGCAGAT GGAGCCCTGG CCATGGAATA GGGAGTGGCC TCTCTGAAGA 3240  
 GTGAGATGAG GGAAGTGGAA GGAGAGCTGG AAAGGAAGGA GCTGGAGTTT GACACGAATA 3300  
 TGGATGCACT ACAGATGCTG ATTACAGAAG CCCAGAAGGT TGATACCAGA GCCAAGAAAG 3360  
 CTGGGGTTAC AATCCAAGAC ACACTCAACA CATTAGACGG CCTCTGCTAT CTGATGAGCC 3420  
 AGCCTCTCAG TGTAGATGAA GAGGGGCTGG TCTTACTGGA GCAGAAGCTT TCCCGAGCCA 3480  
 AGACCCAGCT CAACAGCCAA CTGCGGCCA TGTATGTGCA GCTGGAAGAG AGGCCACGTC 3540  
 AGCAGAGGGG CCACTCCAT TTGCTGGAGA CAAGCATAGA TGGGATTCTG GCTGATGTA 3600  
 AGAATCTGGA GAACATTAGG GACAACCTGC CCCCAGGCTG CTACAATACC CAGGCTCTTG 3660  
 AGCAACAGTG AAGCTGCCAT AAATATTTCT CACTGAGGT TCTTGGGATA CAGATCTCAG 3720  
 GGCTCGGGAG CCATGTGATG TGAGTGGGTG GGTGGGGAC ATTTGAACAT GTTTAATGGG 3780  
 TATGCTCAGG TCACTGACC TGACCCCAT CTGATGCTGG GCAATGAGCG AGATAGCACT GGGTGTGAGA 3840  
 TGCAACATAC CCTTGTCTTC CTGATGCTGG GCAATGAGCG AGATAGCACT GGGTGTGAGA 3900  
 ATGATCAAGG ATCTGGACCC CAAAGATAG ACTGGATGGA AAGACAAACT GCACAGGCAG 3960  
 ATGTTTGCCT CATATCTTGA GTAATGGAG TCCTGGAATT TGGACAAGTG CTGTGGGAT 4020  
 ATAGTCAACT TATCTTTTGA GTAATGTGAC TAAAGGAAA AACTTTGACT TTGCCCAGGC 4080  
 ATGAAATTTT TCCTAATGTC AGAAGCAGAGT GCAACCCAGT CACACTGTGG CCAGTAAAT 4140  
 ACTATTGCTT CATATTGCTT TCTGCAAGCT TCTTGTGAT CAGAGTTCTT CCTACTTACA 4200  
 ACCCAGGGTG TGAACATGTT CTCCATTTC AAGCTGGAAG AAGTGAAGCAG TGTGGAGTG 4260  
 AGGACCTGTA AGGCAGGCC ATTCAAGCT ATGGTGTCTG CTGGTGCCTG CCACCTTCAA 4320  
 GTTCTGGAGC TGGGCATGAC ATCTTTCTT TTAATGATGC CATGGCACT TAGAGATTGC 4380  
 ATTTTATTA AAGCATTTCC TACCAGCAA GCAATGTTG GGAAGTATT TACTTTTTCG 4440  
 GTTTCAAAGT GATAGAAAAG TGTGGCTTGG GCATTGAAAG AGGTAAATTT CTCTAGATT 4500  
 ATTAGTCTTA ATCAATCCT ACTTTGGA CACCAGAAAT GATGCGCATC AATGTATTT 4560  
 ATCTTATTTT CTCAATCCT TCTCTTTTC CTCCACCCAT AATAAGAGAA TGTCTCTACT 4620  
 CACACTTCAG CTGGGTACCA TCCATCCCTT CATTCATCCT TCCATCCATC TTTCCATCCA 4680  
 TTACCTCCAT CCATCCTTCC AACATATAT TATTGATGAC CTACTGTGTG CCAGGGGCTG 4740  
 GTGGGACAGT GGTGACATAG TCTTGGCCT CATAGAGTTG ATTGTCTAGT GAGGAAGACA 4800  
 AGCATTTTAA AAAATAAAT TAAACTTAC AAATTTGTT TGTCAAGT GGTGTTTAT 4860  
 GCAATAACCG CTGTGTTTGC AACCTCTTG CTCACAGAA CATATGTTGC AAGACCTCC 4920  
 CATGGGGCA CTGAGTTTTC GGCAAGGCTG ACAGAGCTCT GGGTTGTGCA CATTTCTTTG 4980  
 CATTCAGCT GTCACTCTGT GCCTTTCTAC AACTGATTGC AACAGACTGT TGAGTTATGA 5040  
 TAACACAGT GGAATTTGCT GGAGGAACCA GAGGCACCTC CACCTTGGCT GGAAGACTA 5100  
 TGGTGTGCC TTGCTTCTGT ATTTCTTGG ATTTCTCTGA AAGTGTTTT AAATAAGAA 5160  
 CAATTGTTAG ATGCC

Seq ID NO: 591 Protein sequence  
 Protein Accession #: NP\_005553.1

80  
 85

1 11 21 31 41 51  
 MPALWLGCLL CFSLLLPAA ATSRREVDC NGKSRQCIFD RELHROTNG FRLCNCNDNT 60  
 DGIHCEKCKV GFYRHRERDR CLPCNCSNG SLARCDNSG RSCCKPVTG ARCDRLPGF 120  
 HMLTDAGCTQ DQRLDLSKCD CDPAGIAGPC DAGRCVCKPA VTGERCDRCR SGYVNLDDGN 180  
 PEGCTQCFY GHASACRSSA EYSVHKITST FHQVDGDKA VQRNGSPAKL QWSQRHQDVP 240  
 SSAQRLLDPY FVAPAKFLGN QVSYGQSLG FDIRVDRGR HPSAHDVILE GAGLRITAPL 300  
 MPLGKTLPCG LTKYTFRLN EHPSNNWSPQ LSYFEYRRL RNLTLALIRA TYGEYSTGYI 360  
 DNVTLISARP VSGAPAPWVE QCICPVGYKG QFCQDCASGY KRDSARLGPP GTCIPCNQGG 420  
 GGACDPDTGD CYSYSDENPDI ECADCPIGFY NDFHDFRSCK PCPCHNPFSC SVMPEETEYV 480

WO 02/086443

PCT/US02/12476

CMCPCPGVTG ARCELADGY FGDPFGEHGP VRPCQPCQCN NNVDPSASGN CDRLTGRCLK 540  
 CIHNTAGIYC DQCKAGYFGD PLAFNPADKC RACNCNPMGS EFVGCSDGT CVCKPGFGGP 600  
 NCEHGAFFSCP ACYNQVKIQM DQFMQQLQRM EALISKAQGG DGVVPDTELE GRMQQAEQAL 660  
 QDILRDAQIS EGASRLSLGQ LAKVRSQENS YQSRLLDLKM TVERVRLGSG QYQNRVRDTH 720  
 RLITQMLSL AESASLGNIT NIPASDHYVG PNGFKSLAQE ATRLAESHVE SASNMQLTR 780  
 ETEDYSKQAL SLVRKALHSG VSGSGSPDG AVVQGLVEKL ETKSLAQQL TREATQABIE 840  
 ADRSYQHSRLR LLDVSRLQG VSDQSPQVEE AKRIKQKADS LSTLVTRHMD EFKRTQKNLG 900  
 NWKEEAQQLL QNGKSGREKS DQLLSRANLA KSRAGEALSM GNATFYEVES ILKNLREFDL 960  
 QVONRKABAE EAMKRLSYIS QKVDASDKT QQAERALGSA AADAQRAKNG AGEALEISSE 1020  
 IEQSIGSLNL EAMVTADGAL AMEKGLASLK SEMREVEGEL ERKELEFDN MDAVQMVITE 1080  
 AQKVDTRAKN AGVTIQDTLN TLDGLLHMD QPLSVDEEGL VLEQLKLSRA KTQINSQLRP 1140  
 NMSELEERAR QQRGHLLELE TSIDGILADV RNLENIRDNL PPGCYNTQAL EQQ

Seq ID NO: 592 DNA sequence  
 Nucleic Acid Accession #: AP101051.1  
 Coding sequence: 221.856

1 11 21 31 41 51  
 20 GAGCAACCTC AGCTTCTAGT ATCCAGACTC CAGCGCCGCC CCGGCGCGCG ACCCCAACCC 60  
 CGACCCAGAG CTCTCTCCAGC GGCGGCGCAG CGAGCAGGGC TCCCCGCTT AACTTCTTCC 120  
 GCGGGGCCCA GCCACCTTCG GGAGTCCGGG TTGCCCACTC GCAAACCTCT CGCCTTCTGC 180  
 ACCTGCCACC CCTGAGCCAG CGCGGCGGCC CGAGCGAGTC ATGCCCAAG CGGGCTGCA 240  
 25 GCTGTTGGGC TTCAATCTCG CCTTCTGGG ATGGATCGGC GCCATCGTCA GCACCTGCCCT 300  
 CGCCCAAGTG AGGATTACT CCTATGCCG CGACAACATC GTGACCGCCC AGGCCATGTA 360  
 CGAGGGGCTG TGGATGTCTT CGGTGTGCA GAGCACCGGG CAGATCCAGT GCAAAGTCTT 420  
 TGACTCCTTG CTGAATCTGA CGACACATT GCAAGCAACC CGTGCCTTGA TGGTGGTTGG 480  
 CATCTCTCTG GGAATGATAG CAATCTTTGT GGCCACCGTT GGCATGAAGT GTATGAAGTG 540  
 30 TTGGAAGAC GATGAGGTGC AGAAGATGAG GATGGCTGTC ATTGGGGGTG CGATATTCTT 600  
 TCTTGCAGT CTGGCTATTT TAGTGTCCAC AGCATGGTAT GGCATAGAA TCGTTCAAGA 660  
 ATTCTATGAC CCTATGACCC CAGTCAATGC CAGGTACGAA TTGTGTCAGG CTCTCTTCC 720  
 TGGCTGGGCT GCTGCTTCTC TCTGCTTCT GGGAGGTGCC CTACTTGTCT GTTCTGTCTC 780  
 CGGAAAAACA ACCTCTTACC CAACACCAAG GCCCTATCCA AAACCTGCAC CTTCCAGCGG 840  
 35 GAAAGACTAC GTGTGACACA GAGGCAAAAG GAGAAAATCA TGTGAAACA AACCGAAAAT 900  
 GGACATTGAG ATACTATCAT TAACATTAGG ACCTTAGAAT TTTGGGTATT GTAATCTGAA 960  
 GTATGATATT ATACAAACAA CAACAAACAA AAAAAACCAT GTGTAAATAT ACTAGTGTCT 1020  
 AAACATGGCT TAATCTTATT TTATCTTCTT TCCTCAATAT AGGAGGGAAG ATTTTACCAT 1080  
 TGTATTACTT GCTTCCCATT GAGTAATCAT ACTCAATGG GGAAGGGGT GCTCCTTAAA 1140  
 40 TATATATAGA TATGATATA TACATGTTT TCTATTAAAA ATAGACAGTA AATACTATT 1200  
 CTCATTATGT TGATACTAGC ATACTTAAAA TATCTTAAA ATAGGTAAAT GTATTTAATT 1260  
 CCATATTGAT AGAGATGTTT ATTGGTATAT TTTCTTTTTC GTCTTATAT ACATATGTAA 1320  
 CAGTCAAAAT TCATTTACTC TTCTTCAAT GCTTTGGGTG CCTTGTCCAC AAGACCTAGC 1380  
 CTAATTTACC AAGGATGAAT TCTTTCAATT CTTCTGCGT GCCCTTTTCA TATACTTATT 1440  
 45 TTATTTTITA CCAATATCTT ATAGCACTTG CATCGTTATT AAGCCCTTAT TTGTTTGTG 1500  
 TTTCTTGTGT CTCTATCTCC TGAATCTAAC ACATTTTATA GCCTACATTT TAGTTTCTAA 1560  
 AGCCCAAGAA AATTATTATC AATCTAGAAC TTTGGAGGCA AATCTTCTG CATGACCAA 1620  
 GTGATAAATC CCTGTTGACC TTCCCAACA ATCCCTGTAC TCTGACCAT AGCACTCTTG 1680  
 TTTGCTTTGA AATATTGTTT CCAATTGAGT AGCTGCATGC TGTTCGCCCA GGTGTGTGTA 1740  
 50 CACAACCTTA TTGATTGAAT TTTTAAGCTA CTTATTCTAT GTTTTATATC CCCTTAAACT 1800  
 ACCTTTTGTG TCCCATCTCC TTAATGTAT TGTTTTCCCA AGTGTAAITA TCATGCGTTT 1860  
 TATATCTTCC TAATAAGGTG TGGTCTGTTT GTCTGAACAA AGTGCTAGAC TTTCTGGAGT 1920  
 GATAATCTGG TGACAAATAT TCTCTCTGTA GCTGTAAACA AGTCACCTAA TCTTCTTACC 1980  
 TCTTTTCTCT ATCTGCCAAA TTGAGATAAT GATACTTAAC CAGTTAGAAG AGGTAGTGTG 2040  
 55 AATATTAAAT AGTTTATATT ACTCTCATTC TTTGAACATG AACTATGCCT ATGTAGTGTG 2100  
 TTTATTGTCT CAGCTGGCTG AGACACTGAA GAAGTCACTG AACAAAACT ACACAGTAC 2160  
 CTTCAATGTA TTCACTGCCCT TCCTCTCTCT ACCAGTCTAT TTCCACTGAA CAAACCTAC 2220  
 ACACATACTT TCATGTGGTT CAGTGCCCTC CTCTCTCTAC CAGTCTATT CCACCTGAAC 2280  
 AAACCTAGCG ACATACCTTC ATGTGGCTCA GTGCCCTTCT CTCTCTACCA GTCTATTTC 2340  
 60 ATTCTTTCAG CTGTGTCTGA CATGTTTGTG CTCTGTTCCT TTTTAAACAC TGCTCTTACT 2400  
 TTTCCAGTCT GTACAGAATG CTATTCTACT TGAGCAAGAT GATGATGGA AAGGGTGTG 2460  
 GCACTGGTGT CTGGAGACCT GGATTTGAGT CTGGGTGCTA TCAATCACCG TCTGTGTTG 2520  
 AGCAAGGCAAT TTGGCTGTCT TAAGCTTATT GCTTCATCTG TAAGCGGTGG TTTGTAATTC 2580  
 CTGATCTTCC CACCTCACAG TGATGTTGTG GGGATCCAGT GAGATAGAA ACATGTAAGT 2640  
 65 GTGTTTGTGT AATTGTAAGA GTGCTATACT AAGGGAAAGA ATTGAGGAAT TAACTGCATA 2700  
 CGTTTGTGTG TTGCTTTTCA AATGTTTGAA AATAAAAAA TGTAAAGAAA TGGTTTCTT 2760  
 GCCTTAACCA GTCTCTCAAG TGATGAGACA GTGAAGTAAA ATTGAGTGCA CTAACGAAT 2820  
 AAGATTCTGA GGAAGTCTTA TCTTCTGCAG TGAGTATGGC CCAATGCTTT CTGTGGCTAA 2880  
 ACAGATGTAA TGGGAAGAAA TAAAGCCTTA CGTGTGGTGA AATCCAACAG CAAGGGAGAT 2940  
 70 TTTTGAATCA TAATAACTCA TAAGGTGCTA TCTGTTCACT GATGCCCTCA GAGCTCTTGC 3000  
 TGTAGCTGG CAGCTGACGC TGCTAGGATA GTTAGTTTGG AAATGCTACT TCATAATAAA 3060  
 CTACACAAGG AAGTCAAGCC ACGGTGCTCT ATGAGGAATT GACCTAATA AATTTTAGTG 3120  
 TGCCTTCCAA ACCTGAGAA ATATGCTTTT GGAAGTTAAA ATTAAATGG CTTTGGCCAC 3180  
 ATACATAGAT CTTCATGATG TGTAGTGTGA ATTCCATGTG GATATCAGT ACCAACATT 3240  
 75 AAAAAAAT TTTATGCCCC AAAATGACCA ACGAAATGT TACAATAGAA TTTATCCAAT 3300  
 TTTGATCTTT TTATATTCTT CTACCAACC TGGAAACAGA CCAATAGACA TTTGGGGTT 3360  
 TTATAATGGG AATTTGTATA AAGCATTACT CTTTTTCAAT AAATGTTTT TTAATTTAAA 3420  
 AAAAGGAAAA AAAAAAAA AAA

Seq ID NO: 593 Protein sequence  
 Protein Accession #: AAD16433.1

1 11 21 31 41 51  
 85 MANAGLQLLG FILAFLGWIG AIVSTALPQW RIYSYAGDNI VTAQAMYEGE WMSCVSQSTG 60  
 QIQCKVDSL LNLSSLTQAT RALMVVGILL GVIAIPVATV GMCKMKLED DEVQKMRMAV 120  
 IGGAIPLLAG LAILVATAWY GNRIVQEFYD PMTPVNARYE FGQALFTGWA AASLCLLGA 180  
 LLCCSCPRTK TSYPTRPRYP KPRESSGKDY V



Seq ID NO: 594 DNA sequence  
Nucleic Acid Accession #: NM\_006180.1  
Coding sequence: 352..2820

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70

1	11	21	31	41	51	
CCCCATTGG	CATCTAACAA	GGAATCTGCG	CCCCAGAGAG	TCCCGGACGC	CGCCGGTCGG	60
TGCCCGGCGC	GCCGGGCCAT	GCAGCGACGG	CCCGCGCGGA	GCTCCGAGCA	GCGGTAGCGC	120
CCCCCTGTAA	AGCGGTTCGC	TATGCGGGGA	CCACTGTGAA	CCCTGCGGCC	TGCCGGAACA	180
CTCTTCGCTC	CGGACCAGCT	CAGCCTCTGA	TAAGCTGGAC	TGGGACGCCC	CGCAACAAGC	240
ACCGAGGAGT	TAAGAGAGCC	GCAAGCGCAG	GGAAGGCCTC	CCCGCACGGG	TGGGGGAAAG	300
CGCCCGGTGC	AGCGCGGGGA	CAGGCACTCG	GGCTGGCACT	GGCTGCTAGG	GATGTCGTCC	360
TGGATAAGGT	GGCATGAGCC	CGCCATGGCG	CGGCTCTGGG	GCTTCTGCTG	GCTGGTTGTG	420
GGCTTCTGGA	GGCGCGCTTT	CGCCTGTCCC	ACGTCCTGCA	AATGCAGTGC	CTCTCGGATC	480
TGGTGACAGG	ACCCTTCTCC	TGGCATGGTG	GCATTTCGGA	GATTGGAGCC	TAAACAGTGA	540
GATCTTGAGA	ACATCAACGA	AATTTTCATC	GCAAAACAGA	AAAGGTTAGA	AATCATCAAC	600
GAAGATGATG	TTGAAGCTTA	TGTGGGAGTG	AGAAATCTGA	CAATTGTGGA	TTCTGGATTA	660
AAATTTGTGG	CTCATAAAGC	ATTCTTGAAA	AACAGCAACC	TGCAGCACAT	CAATTTTACC	720
CGAAACAAAC	TGACGAGTTT	GTCTAGGAAA	CATTTCCTCG	ACCTTGACCT	GTCTGAACTG	780
ATCCTGTGGG	GCAATCCATT	TACATGCTCC	TGTGACATTA	TGTGGATCAA	GACTCTCCAA	840
GAGGCTAAAT	CCAGTCCAGA	CACTCAGGAT	TTGTACTGCC	TGAATGAAAG	CAGCAAGAAT	900
ATTCCTCTGG	CAAACTTGCA	GATACCCAAT	TGTGGTTTGC	CATCTGCAAA	TCTGGCCGCA	960
CCTAACCTCA	CTGTGGAGGA	AGGAAAGTCT	ATCACATTAT	CCTGTAGTGT	GGCAGGTGAT	1020
CGGTTTCTTA	ATATGTATTG	GGATGTTGGT	AACCTGGTTT	CCAAACATAT	GAATGAAACA	1080
AGCCACACAC	AGGGCTCCTT	AAGGATAACT	AACATTTTCA	CGATGACAG	TGGGAAGCAG	1140
ATCTCTTGTT	TGGCGGAAAA	TCTTGTAGGA	GAAGATCAAG	ATTCTGTCAA	CCTCACTGTG	1200
CATTTTGCAC	CAACTATCAC	ATTTCTCGAA	TCTCCAACCT	CAGACCCACA	CTGGTGCATT	1260
CCATTCTAGT	TGAAAGGCAG	CCCCAAACCA	CGGCTTCAGT	GGTTCTATAA	CGGGGCAATA	1320
TTGAATGAGT	CCAAATACAT	CTGTACTAAA	ATACATGTTA	CCAATCACAC	GGAGTACCAC	1380
GGCTGCCTCC	AGCTGGATAA	TCCCACTCAC	ATGAACAATG	GGGACTACAC	TCTAATAGCC	1440
AGAATGAGT	ATGGGAAGGA	ATTCTGCTC	ACTTCAATGG	CTGGCTGGA		1500
ATTGACGAGT	GTGCAAAACC	AAATATCTCT	GATGTAATTT	ATGAAGATTA	TGGAACATGCA	1560
GGCAATGACA	TGGGGGACAC	CACGAACAGA	AGTAATGAAA	TCCCTTCCAC	AGACCTCACT	1620
GATAAAACCG	GTCCGGGAACA	TCTCTGGTCT	TATGCTGTGG	TGGTGAATGC	GTCTGTGGTG	1680
GGATTTTGCC	TTTTGGTAAT	GCTGTTTCTG	CTTAAGTTGG	CAAGACATCT	CAAGTTTGGC	1740
ATGAAAGGCC	CAGCCTCCGT	TATCAGCAAT	GATGATGACT	CTGCCAGCCC	ACTCCATCAC	1800
ATCTCCAATG	GGAGTAACAC	TCCATCTCTT	TCCGAAAGTG	GCCCAGATGC	TGTCATTATT	1860
GGATGACCA	AGATCCCTGT	CATTGAAAAT	CCCCAGTACT	TTGGCATCAC	CAACAGTCAG	1920
CTCAAGCCAG	ACACATTTGT	TACGACATC	AAGCGACATA	ACATTGTTCT	GAAAAGGGAG	1980
CTAGGCGAAG	GAGCCTTTGG	AAAAGTGTTC	CTAGCTGAAT	GCTATAACCT	CTGTCTGAG	2040
CAGGACAGA	TCTTGGTGGC	AGTGAAGACC	CTGAAGGATG	CCAGTGACAA	TGCAACGCAAG	2100
GACTTCCACC	GTGAGGCCGA	GCTCCTGACC	AACCTCCAGC	ATGAGCACAT	CGTCAAGTTC	2160
TATGGCGTCT	GGGTGGAGGG	CGACCCCTCT	ATCATGCTCT	TTGAGTACAT	GAAGCATGGG	2220
GACCTCAACA	AGTTCTCTAG	GGCACAAGCG	CCTGATGCCG	TGCTGATGGC	TGAGGGCAAC	2280
CGCCCCACGG	AACGAGCGCA	GTGCGAGATG	CTGCATATAG	CCAGCAGATG	CGCCGCGGGC	2340
ATGGTCTTGC	TGGCGTCCCA	GCACCTCTGT	CACCGGAGTT	TGGCCACAG	GAAGTGCCTG	2400
GTGCGGGAGA	ACTTGTCTGT	GAAAATCGGG	GACTTTGGGA	TGTCCCGGGA	CGTGTACAGC	2460
ACTGACTACT	ACAGGGTCGG	TGGCCACACA	ATGCTGCCCA	TTGCTGGAT	GCCTCCAGAG	2520
AGCATCATGT	ACAGGAAATT	CACGACGGAA	AGCGACGCTT	GGAGCCTGGG	GGTCTGTGTT	2580
TGGAGAGATT	TCACTATGAG	CAAAACAGCC	TGGTACCAGC	TGTCAACAA	TGAGGTGATA	2640
GAGTGTATCA	CTCAGGGCCG	AGTCTCTGAG	CGACCCCGCA	CGTCCGCCCA	GGAGGTGTAT	2700
GAGCTGATGC	TGGGTTGCTG	GCAGCGAGAG	CCCCACATGA	GGAAGAACAT	CAAGGGCATC	2760
CATACCTCTC	TTCAAGACTT	GGCCAAAGCA	TCTCCGCTCT	ACCTGGACAT	TCTAGGCTAG	2820
GGCCCTTTTC	CCGACACCGA	TCTTCCCAAA	CGTACTCCTC	AGACGGGCTG	AGAGGATGAA	2880
CATCTTTTAA	CTGCGCTGGG	AGGCCAACAA	GCTGCTCTCC	TTCACTCTGA	CAGTATTAAAC	2940
ATCAAAGACT	CGAGAGAGCT	CTCGAGGGAA	GCAGTGTGTA	CTTCTCTATC	CATAGACACA	3000
GTATTGACTT	CTTTTGGCCA	TTATCTCTTT	CTCTCTTTCC	ATCTCCCTTG	GTGTGTCCTT	3060
TTTCTTTTIT	TAAATTTTCT	TTTTCTCTCT	TTTTTCTGTC	TTCCCTGCTT	CAOGATTCTT	3120
ACCCTTTCTT	TTGAATCAAT	CTGGCTCTCG	CATTACTATT	AACCTCTGCAT	AGACAAGGCG	3180
CTTAACAAAC	GTAATTTGTT	ATATCAGCAG	ACACTCCAGT	TTGCCACCA	CAACTAACAA	3240
TGCTTGTGTT	TATTCCTGCC	TTTGATGTGG	ATGAAAAAAA	GGGAAAAACA	ATATTTCACT	3300
TAAACTTTGT	CACCTCTGCT	GTACAGATAT	CGAGAGTTTC	TATGGATTCA	CTTCTATTTA	3360
TTTATTATTA	TTACTGTTCT	TATTGTTTTT	GGATGGCTTA	AGCCTGTGTA	TAAAAAGAA	3420
AACTGTGTGT	CAATCTGTGA	AGCCTTTATC	TATGGGAGAT	TAAAAACAGA	GAGAAAGAAG	3480
ATTTATTATG	AACGCAATA	TGGGAGGAAC	AAAGACAACC	ACTGGGATCA	GCTGGTGTCA	3540
GTCCCTACTT	AGGAAATACT	CAGCAACTGT	TAGCTGGGAA	GAATGTATTC	GGCACCTTCC	3600
CCTGAGGACC	TTTCTGAGGA	GTAAGAAAGAC	TACTGGCCTC	TGTGCCATGG	ATGATTCTTT	3660
TCCCATCACC	AGAAATGATA	GCTGTCAGTA	GAGAGCAAAG	ATGGCTT		

Seq ID NO: 595 Protein sequence  
Protein Accession #: NP\_006171.1

75  
80  
85

1	11	21	31	41	51	
MSSWIRWEGP	AMARLWGFCH	LUVGFWRAAF	ACPTSCCKSA	SRIWCSDPSP	GIVAFPRLEP	60
NSVDPENITE	IFIANQKRLK	IINEDDVEAY	VGLRNLITVD	SLGLFVAHKA	FLKNSNLQHI	120
NFTRNKLTL	SRKHFRHLDL	SELILVGNPF	TCSCDIMWIK	TLQEAQSSPD	TQDLYCLINES	180
SKNIPLANLQ	IPNCGLPFSN	LAAFNLTVEE	GKSTILSCSV	AGDPVPNNMY	DVGNLVSKHM	240
WTSHTQGLSL	RITNISDDDS	GKQISCVAE	NLVGEDQDSVN	LTVHPAPTIT	FLESPTSDDH	300
WCIPFTVKGN	PKPALQWFFN	GAILNESKYI	CTKLHVTNHT	EYHGCLQLDN	PTHMNGDYT	360
LIARNEYGKD	EKQISAHFMG	WFGIDDGANP	NYPDVIYEDY	GTAANDIGDT	TNRSEIFST	420
DVTDKTRGHE	LSVYAVVVIA	SVVGFCLLVM	LFLKLKLARH	KFGMKGPASV	ISNDDDSASP	480
LHLISNGSNT	PSSSEGGPDA	VIIGMTKIPV	IENPYQFGIT	NSQLKPDFTV	QHIKREHIVL	540
KRELGEAGFG	KVFLAECYNL	CPEQDKILVA	VKTLKDASDN	ARKDPFHRAE	LLTNLQHEHI	600
VKPYGVCEVG	DELIMVFEYM	KRGDLNKPLR	AHGFDAVIMA	EGNPPELTQ	SQMLHIAQOI	660
AAGMYYLASQ	HFVHRDLATR	NCLVGENLLV	KIGDFGMSRD	VYSTDYRVVG	GHTMLPIRMW	720



PPESIMYRKP TTESDVNSLG VVLWEIFTYG KQPMYQLSNN EVIECITQGR VLQRPRTPCQ 780  
 EYVELMLGCH QREPHMRKNI KGIHTLLQNL AKASPVYLDI LG

Seq ID NO: 596 DNA sequence  
 Nucleic Acid Accession #: AF410899  
 Coding sequence: 483..2999

	1	11	21	31	41	51	
10	GGGAGCAGGA	GCCTCGCTGG	CTGCTTCGCT	CGCGCTCTAC	GCGCTCAGTC	CCCGCGGGTA	60
	GCAGGAGCCT	GGACCCAGGC	GCCGCGCGCG	GGCGTGAGGC	GCGGAGCCCC	GGCCTCGAGG	120
	TGCATACCGG	ACCCCATTC	GCATCTAACA	AGGAATCTGC	GCCCCAGAGA	GTCCCGGACG	180
	CCGCGCGTGG	GTGCGCGCGG	CGCGCGCGCG	TGCAGCGACG	GCCGCGCGCG	AGCTCGAGC	240
15	AGCGGTAGCG	CCCCCTGTGA	AAGCGGTTGG	CTATGCGGGG	ACCACTGTGA	ACCTGCGCGC	300
	CTGCGCGAAC	ACTCTTCGCT	CGGACCCAGC	TCAGCCTCTG	ATAAGCTGGA	CTCGGCACGC	360
	CCGCAACAAG	CACCGAGGAG	TTAAGAGAGC	CGCAAGCGCA	GGGAAGGCCT	CCCCGACGG	420
	GTGCGGGAAG	GCGGCGGTGG	CAGCGCGGGG	ACAGGCACCT	GGGCTGGCAC	TGGCTGCTAG	480
	GGATGTCGTC	CTGATAAAGG	TGGCATGGAC	CCGCCATGGC	GCGCTCTGGG	GGCTTCTGCT	540
20	GGCTGGTGTG	GGGCTTCTGG	AGGGCGGCTT	TCGCCTGTCC	CAGCTCCTGC	AAATGCAGTG	600
	CCTCTCGGAT	CTGGTGCAGC	GACCCCTTCT	CTGGCATCGT	GGCATTTCCG	AGATTGGAGC	660
	CTAACAGTGT	AGATCTCTAG	AACATCACCG	AAATTTTCAT	CGCAAAACAG	AAAAGGTTAG	720
	AAATCATCAA	CGAAGATGAT	GTGGAAGCTT	ATGTGGGACT	GAGAAATCTG	ACAATTGTGG	780
	ATTCTGGATT	AAAATTTGTG	GCTCATAAAG	CATTTCGTAA	AAACAGCAAC	CTGCAGCACA	840
25	TCAATTTTAC	CGAAAACAAA	CTGACGAGTT	TGTCTAGGAA	ACATTTCCGT	CACCTTGACT	900
	TGTCGTAAGT	GATCCTGGTG	GGCAATCCAT	TTACATGCTC	CTGTGACATT	ATGTGGATCA	960
	AGACTCTCCA	GAGCTCTAAA	TCCAGTCCAG	ACACTCAGGA	TTTGTAAGTC	CTGAATGAAA	1020
	GCAGCAAGAA	TATTCCTCTG	GCAAACTCTG	AGATACCCAA	TTGTGGTTTG	CCATCTGCAA	1080
	ATCTGGCGCG	ACCTAACCTC	ACTGTGGAGG	AAGGAAAGTC	TATCACATTA	TCCTGTAGTG	1140
30	TGGCAGGTGA	TCGGGTTCCCT	AATATGTATT	GGGATGTTGG	TAACTCGGTT	TCCAAACATA	1200
	TGAATGAAC	AAGCCACACA	CAGGGCTCCT	TAAGGATAAC	TAACATTTCA	TCCGATGACA	1260
	GTGGGAGACA	GATCTCTTGT	GTGGCGGAAA	ATCTTGTAGG	AGAAGATCAA	GATTCTGTCA	1320
	ACCTCACTGT	GCAATTTGCA	CCAATCTACA	CATTCTCGGA	ATCTCCAAAC	TCAGACCAAC	1380
	ACTGGTGCAAT	TCCATTCACT	GTGAAAGGCA	ACCCCAACCC	AGCGCTTCAG	TGGTTCTATA	1440
35	ACGGGGCAAT	ATTGAATGAG	TCCAAATACA	TCTGTACTAA	AATACATGTT	ACCAATCACA	1500
	CGGAGTACCA	CGGCTGCCTC	CAGCTGGATA	ATCCCACTCA	CATGAACAAT	GGGGACTACA	1560
	CTCTAATAGC	CAAGAATGAG	TATGGGAAGG	ATGAGAAACA	GATTTCTGCT	CACCTTCATG	1620
	GCTGGCTGGG	AATTGACGAT	GGTGCAAAAC	CAAATTTATC	TGATGTAATT	TATGAAGATT	1680
	ATGGAATGCG	AGCGAATGAC	ATCGGGGACA	CCACGAACAG	AAGTAATGAA	ATCCCTTCCA	1740
40	CAGACGTCAC	TGATAAAACC	GGTCGGGAAC	ATCTCTCGGT	CTATGCTGTG	GTGGTGATTG	1800
	CGTCTGTGGT	GGGATTTTGC	CTTTTGGTAA	TGCTGTTTCT	GCTTAAGTTG	GCAAGACACT	1860
	CCAAGTTTGG	CAAGAAGATG	TTCTCATGGT	TTGGATTGGG	GAAAGTAAAA	TCAGAGACAG	1920
	GTGTTGGCCC	AGCCTCGGTT	ATCAGCAATG	ATGATGACTC	TGCCAGCCCA	CTCCATCACA	1980
	TCTCAATGCG	GAGTAACACT	CCATCTTCTT	CGAAGGTGGG	CCAGATGCT	GTCAATTATG	2040
45	GAATGACCAA	GATCCCTGTC	ATTGAAAATC	CCCAGTACTT	TGGCATCACC	AACAGTCAGC	2100
	TCAAGCCAGA	CACATTGTGT	CAGCACATCA	AGCGACATAA	CATTGTTCTG	AAAAGGGAGC	2160
	TAGCGGAAGG	AGCCTTTGGA	AAAGTGTTC	TAGCTGAATG	CTATAACCTC	TGTCCTGAGC	2220
	AGGACAGAGT	CTTGGTGGCA	GTGAAGACCC	TGAAGGATGC	CAGTGACAA	GCACGCAAGG	2280
	ACTTCCACCG	TGAGGCGGAG	CTCCTGACCA	ACCTCCAGCA	TGAGCACATC	GTCAAGTTCT	2340
50	ATGGCGTCTG	CGTGGAGGGC	GACCCCTCTA	TCTGTTCTTT	TGAGTACATG	AAGCATGGGG	2400
	ACCTCAACAA	GTTCCTCAGG	GCACACGGCC	CTGATGCCGT	GCTGATGGCT	GAGGGCAACC	2460
	CGCCACCGGA	ACTGACGCGG	TCCGAGATGC	TGCATATAGC	CAGACAGATC	GCGCGGGGCA	2520
	TGGTCTACCT	GGCGTCCGAG	CACCTCTGTC	ACCGCGATTT	GGCCACGAGG	AACTGCCTGG	2580
	TCGGGGAGAA	CTTGTCTGTT	AAAATCGGGG	ACTTTGGGAT	GTCCCGGGAC	GTGTACAGCA	2640
55	CTGACTACTA	CAGGTCGGTG	GGCCACACAA	TGCTGCCCAT	TGCTGGATG	CCTCCAGAGA	2700
	GCACTCATGA	CAGGAATTC	ACGACGGAAG	GCGACGCTCG	GAGCCTGGGG	GTGCTGTTGT	2760
	GGGAGATTTT	CACCTATGGC	AAACAGCCCT	GGTACCAGCT	GTCAAAACAT	GAGGTGATAG	2820
	AGTGATACAC	TCAGGGCGGA	GCTCTGCAGC	GACCCCGCAG	GTGCCCCCAG	GAGGTGTATG	2880
	AGCTGATGCT	GGGGTCTGTT	CAGCGAGAGC	CCCACATGAG	GAAGAACATC	AAGGGCATCC	2940
60	ATACCCCTCT	TCAGAACTTG	GCCAAGGCAT	CTCCGGCTTA	CCTGGACATT	CTAGGCTAGG	3000
	GCCCTTTTCC	CCAGACCGAT	CCTTCCCAAC	GTACTCTCTA	GACGGGCTGA	GAGGATGAAC	3060
	ATCTTTTAAC	TGCGCTGGA	GGCCACCAAG	CTGCTCTCCT	TCACTCTGAC	AGTATTAACA	3120
	TCAAAGACTC	CGAAGAGCTC	TGAGGGGAAG	CAGTGTGTAC	TTCTTCATCC	ATAGACACAG	3180
	TATTGACTTC	TTTTTGGCAT	TATCTCTTTC	TCTCTTCCA	TCTCCCTTGG	TTGTTCTTTT	3240
65	TTCTTTTCTT	AAATTTTCTT	TTTCTCTTTC	TTTTTCTGCT	TCCCTGCTTC	ACGATTCTTA	3300
	CCCTTTCTTT	TGAATCAATC	TGGCTCTGTC	ATTACTATTA	ACTCTGCATA	GACAAAGGCC	3360
	TTAACAACCG	TAATTTGTTA	TATCAGCAGA	CACCTCCAGT	TGCCCAACAC	AACTAACAAT	3420
	GCCTTGTGTT	ATTCTCGCCT	TTGATGTGGA	TGAATAAAGG	GGAAACAAA	TATTTCACTT	3480
	AAACTTGTTC	ACTTCTGCTG	TACAGATATC	GAGAGTTTCT	ATGGATTAC	TTCTATTTAT	3540
70	TTATTATTAT	TACTGTTCTT	ATTGTTTGTG	GATGGCTTAA	GCCTGTGTAT	AAAAAGAAAA	3600
	ACTTGTGTTT	AATCTGTGAA	GCCTTTATCT	ATGGGAGATT	AAAACCAAGG	AGAAAGAGAA	3660
	TTTATTATGA	ACCGCAATAT	GGGAGGAACA	AAGACACCA	CTGGGATCAG	CTGGTGTCA	3720
	TCCTACTTAA	GGAAATACTC	AGCAACTGTT	AGCTGGGAAG	AATGTATTGG	GCACCTTCCC	3780
	CTGAGGACCT	TTCTGAGGAG	TAAAAAGACT	ACTGGCCTCT	GTGCCATGGA	TGATTCTTTT	3840
75	CCCATCACA	GAAATGATAG	CGTGCACTAG	AGAGCAAGA	TGGCTTCCGT	GAGACACAAG	3900
	ATGGCGCATA	GTGTGCTCGG	ACACAGTTT	GTCTTCTGAG	GTGTGATGA	TAGCACTGGT	3960
	TTGTTTCTCA	AGCGCTATCC	ACAGAACCTT	TGTCAACTTC	AGTTGAAAG	AGGTGGATTTC	4020
	ATGTCACAGG	CTCATTTCGG	GGTCAGGTGG	GAAAGCC			

Seq ID NO: 597 Protein sequence  
 Protein Accession #: AAL67965.1

	1	11	21	31	41	51	
85	MSSWIRWEP	AMARLWGF	LVVGFWRAP	ACPTSCCKSA	SRIWCSDPSP	GIVAFPRLEP	60
	NSVDPENITE	IFIANQKRL	IINEDDVEAY	VGLRNLITVD	SLGKPVAKHA	FLKNSNLQHI	120
	NFTNKLTL	SRKHFRHLD	SELILVGNFP	TCSCDIMWIK	TLQEAQSPD	TQDLYCLNES	180
	SKNIPLANLQ	IPNCGLPAN	LAAPNLVVEE	GKSITLSCSV	AGDPVFNMYW	DVGNLVSKHM	240

NETSHYQSL RITNISSDDS GKQISCAEN LVGEDQDSVN LTVHFAPTIT FLESPTSDDH 300  
 WCIPTFVKGK PKPALQWFYN GAILNESKYI CTKIHVTNHT EYHGLQLDN PTHMNGDYT 360  
 LIANNEYGKD EKQISAHFMG WPGIDDGANP NYPDVIYEDY GTAANDIGDT TNRSNEIPST 420  
 DVIDTKGREH LSVYAVVVIA SVVGFCLVM LFLKLARHS KFGMKDPSWF GFGKVKSRQG 480  
 VGPASVISND DDSASPLHHI SNGSNTSPSS EGGPDVAVIG MTKIPVIEHP QVFGITNSQL 540  
 KPDTFVQHIK RHNIVLKREL GSGAFGRVFL AECYNLCPEQ DKILVAVKTL KDASDNARKD 600  
 FHREAEILLN LQHEHIVKFI GVCVEGDPLI MVFEYMKHGD LNKFLRAHGP DAVLMAEGNP 660  
 PTELTQSQML HIAQQAAGM VYLAHQHFVH RDLATRNCLV GENLLVKIGD FGMSRDVYST 720  
 DYYRVGSHTM LPIRMPPES IMYRKFTTES DVNSLGVVLN EIFTYKQPW YQLSNNEVIE 780  
 CITQGRVLQR PRICPQEVYE LMLGCWQREP HMRKNIKGIH TLLQNLAKAS PVYLDILG

Seq ID NO: 598 DNA sequence  
 Nucleic Acid Accession #: AB052906  
 Coding sequence: 74..814

1 11 21 31 41 51  
 AAAACCTTGA GGTGATTCAT CTCACAGGCT CTCCTTCCAT CAAGTCTCTC CTCCCTAGCG 60  
 CTCCTGGTCC TTAATGGCAG CAGCCGCGGC TACCAAGATC CTTCTGTGCC TCCCGCTTCT 120  
 GCTCCTGCTG TCCGGCTGGT CCGGGCTGGG GCGAGCCGAC CCTCACTCTC TTGCTATGA 180  
 CATCACCGTC ATCCCTAAGT TCAGACCTGG ACCACGCTGG TGTGGGTTC AAGGCCAGGT 240  
 GGATGAAAAG ACTTTTCTTC ACTATGACTG TGGCAACAAG ACAGTCACAC CTGTCAGTCC 300  
 CCTGGGGAAG AAATAAATG TCACACCGGC CTGGAAGAAG CAGAACCCAG TACTGAGAGA 360  
 GGTGGTGGAC ATACTTACAG AGCAACTGGC TGACATTCAG CTGGAGAATT ACACACCCAA 420  
 GGAACCCCTC ACCCTCAGG CCAGGATGTC TTGTGAGCAG AAAGCTGAAG GACACAGCAG 480  
 TGGATCTTGG CAGTTCAGTT TCGATGGGCA GATCTTCTCT CTCTTTGACT CAGAGAAGAG 540  
 AATGTGGACA ACGGTTTATC CTGGAGCCAG AAAGATGAAA GAAAAGTGGG AGAATGACAA 600  
 GGTGTGGGCC ATGTCTTCCC ATTACTTCTC AATGGGAGAC TGTATAGGAT GGCTTGAGGA 660  
 CTTCTTGATG GGCATGGACA GCACCTGGA GCCAAGTGCA GGAGCACCAC TCGCCATGTC 720  
 CTCAGGCACA ACCCACTCA GGGCCACAGC CACCACCTCT ATCCTTTGCT GCCTCCTCAT 780  
 CATCCTCCCC TGCTTCTATC TCCCTGGCAT CTGAGGAGAG TCCTTTAGAG TGACAGGTGA 840  
 AAGCTGATAC CAAAAGGCTC CTGTGAGCAC GGTCTTGATC AAACCTGCCC TTCTGTCTGG 900  
 CCAGCTGCCC ACGACCTACG GTGTATGTCC AGTGGGCTCC AGCAGATCAT GATGACATCA 960  
 TGGACCAAT AGCTCATTC A CTGCTTGAT TCCTTTTGCC AACAAATTTA CCAGCAGTTA 1020  
 TACCTAACAT ATTATCAAT TTCTCTTGG TGCTACCTGA TGGAAATCCT GCACCTTAAAG 1080  
 TTCTGGCTGA CTAAACAAGA TATATCATTT TCTTTCTTCT CTTTGTGTTT GGAATCAAA 1140  
 GTACTTCTTT GAATGATGAT CTCCTTCTTG CAAATGATAT TGTGAGTAAA ATAATCACGT 1200  
 TAGACTTCAG ACCTCTGGGG ATCTCTTCCG TGTCTCTGAA GAGAAATTTT AAATTAATTA 1260  
 ATAAGAAAAA ATTTATATTA ATGATTGTTT CCTTTAGTAA TTTATTGTTT TGTACTGATA 1320  
 TTTAAATAAA GAGTTCTATT TCCCAAAAAA AAAAAAAAAA AA

Seq ID NO: 599 Protein sequence  
 Protein Accession #: BAB61048.1

1 11 21 31 41 51  
 MAAAAATKIL LCLPLLLLS GWSRAGRADP HSLCYDITVI PKFRPGPRWC AVQGGVDEKT 60  
 FLHYDOGNKT VTFVSPGLKK LNVTTAWKAQ NPVLRBVVDI LTELRLDIQL ENYTPKEPLT 120  
 LQARMSCEQR AEGHSSGSWQ PSFDGQIFLL FDSEKRMWTT VHPGARKMKE KWENDKVVAM 180  
 SFHYFSMGDC IGWLEDFLMG MDSTLEPSAG APLAMSSGTT QLRATATTLI LCCLLIILPC 240  
 FILPGI

Seq ID NO: 600 DNA sequence  
 Nucleic Acid Accession #: NM\_001898.1  
 Coding sequence: 57..482

1 11 21 31 41 51  
 GGCTCTCACC CTCCTCTCCT GCAGCTCCAG CTTTGTGCTC TGCTCTGAG GAGACCATGG 60  
 CCCAGTATCT GAGTACCCCT CTGCTCTCTG TGGCCACCCCT AGCTGTGGCC CTGGCCTGGA 120  
 GCCCAAGGA GGAGGATAGG ATAATCCCGG GTGGCATCTA TAACGCAGAC CTCATGATG 180  
 AGTGGGTACA GCGTGGCCTT CACTTCGCCA TCAGCAGATA TAACAAGGCC ACCAAAGATG 240  
 ACTACTACAG ACGTCCGCTG CCGGTACTAA GAGCCAGGCA ACAGACCGTT GGGGGGTGA 300  
 ATTACTTCTT CGAGCTAGAG GTGGGCGGCA CCAATGATG CAAGTCCCGA CCCAAGCTTG 360  
 ACACCTGTGC CTTCCATGAA CAGCCAGAAC TGCAGAAGAA ACAGTTGTGC TCTTTCGAGA 420  
 TCTACGAAGT TCCCTGGGAG AACAGAAGGT CCCTGGTGAA ATCCAGGTGT CAAGAATCCT 480  
 AGGGATCTGT GCCAGGCCAT TCGCACCAGC CACCACCCAC TCCACCCCTC TGTAGTGTCT 540  
 CCACCCCTGG ACTGGTGGCC CCCACCTGCG GGGAGGCTCT CCAATGTGCC TGCGCCAAGA 600  
 GACAGACAGA GAAGGCTGCA GGAGTCTCTT GTTGCTCAGC AGGGCGCTCT GCCCTCCCTC 660  
 CTTCTCTCTT GCTTCTAATA GCCCTGGTAC ATGGTACACA CCCCCCACC TCCTGCAATT 720  
 AAACAGTAGC ATCGCC

Seq ID NO: 601 Protein sequence  
 Protein Accession #: NP\_001889.1

1 11 21 31 41 51  
 MAQYLSTLLL LLATLAVALA WSPKEEDRII PGGIYNADLN DEWVQALHF AISEYNKATK 60  
 DDYRRRLRV LRARQTVGG VNYFFDVZVG RTICTKSQPN LDTCAPEHQ ELQKKQLCSF 120  
 EIYEVFWENR RSLVKSRCQE S

Seq ID NO: 602 DNA sequence  
 Nucleic Acid Accession #: NM\_003976.2  
 Coding sequence: 299..961

1 11 21 31 41 51

CTCTGAGCTT CTCTGAGCCT TGTITGCTCA TCTGGAAAA GGGGATTAAA CCAITTTACCT 60  
CATGGAGTTG TGAAGAATA GCTGCAAAGC ACCTAACACA TAGTAAGGTT CCCAGTGCAG 120  
CTACTTCTGC TGGTTGAGT CTAGCTGTGT AGGCCCTCTG TTCTTCACCT GGAGAACTG 180  
GGGTGGCAGG CCGGTCCCC ACAAAAGATA ACTCATCTCT TAATTGCAA GCTGCTCAA 240  
CAGGAGGTGA GGGGAACAGC TCAACATGCG CTGATGGGCG CTCTGGTGT TGATAGAGAT 300  
GGAACCTGGA CTTGAGGCGC TCTCCACGCT GTCCCACTGC CCCTGGCTTA GGGGAGACC 360  
TGCCCTGTGG CCCACCTCGG CCGCTCTGGC TCTGCTGAGC AGCTGCGCAG AGGCTCCCT 420  
GGGCTCGCGG CCCCGCAGCC CTGCCCCCGG CGAAGGCCCC CCGCTGTGTC TGGCGTCCCC 480  
CGCCGCGCAC CTGCGGGGG GACGACGCGC CCGCTGGTGC AGTGAAGAGC CCGGCGGCG 540  
GCCGCGCAGC CCTTCTCGGC CCGCGCCCCG GCGGCTGCA CCCCCATCTG CTCTTCCCCG 600  
GGGGGGCGCG GGGCGCGGG CTGGGGGCGG GGGCAGCGCG GCTCGGGCAG CCGGGGCGCG 660  
GGGCTGCGCG CTGCGCTCGC AGCTGGTGGC GGTGCGGCGG CTCGGCTGGG GCCACCGCTC 720  
CGACGAGCTG GTGCGTTTCC GCTTCTGCGC CCGCTCCTGC CGCGCGCGCG GCTCTCCACA 780  
CGACCTCAGC CTGCGCAGCC TACTGGGCGC CGGGGCGCTG CGACCGCCCC CGGGCTCCCG 840  
GCCGCTCAGC CAGCCCTGCT GCGGACCCAC GCGCTACGAA GCGGTCTCTT TCATGAGCT 900  
CAACAGCACG TGAAGAACCG TGAACGCGCT CTCGCCACCC GCCTGCGGCT GCGTGGGCTG 960  
AGGGCTCGCT CCAAGGCTTT GCAGACTGGA CCTTACCGG TGGCTCTTCC TGCTGGGAC 1020  
CCTCCCGCAG AGTCCCACTA GCCAGCGGCC TCAGCCAGGG ACGAAGGCTT CAAAGCTGAG 1080  
AGGCCCTAC CTGCGGGTGA TGGATATCAT CCCGGAACAG GTGAAGGGAC AACTGACTAG 1140  
CAGCCCGAGA GCGCTCACCC TGCGGATCCC AGCCTAAAAG ACACAGAGA CCTCAGCTAT 1200  
GGAGCCCTTC GGAGCCACTT CTCACAGACT CTGGCACTGG CCAGGCTCGA AACCTGGGAC 1260  
CCCTCCTCTG ATGAACACTA CAGTGGCTGA GGCATCAGCC CCGGCCAGG CCCTGTAGGG 1320  
ACAGCATTTG AAGGACACAT ATTGCACTTG CTTGGTTGAA AGTGCTGTG CTGGAACCTG 1380  
CTGTACTCA CTATGGGAG CTGGCCCC

Seq ID NO: 603 Protein sequence  
Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
RGCLRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180  
RPVSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 604 DNA sequence  
Nucleic Acid Accession #: NM\_057091.1  
Coding sequence: 783..1445

1 11 21 31 41 51  
ACTGGCGCGT GAGAGAAGAA TCGGGTGGAG CAGAGAGCAG CTGCTGCAGG GCAGACAGCC 60  
GGACCCCAA ATCTGCACGT ACCAGCAGTC AGCCGCCCA CGCAGGGACC GGCTTACCCC 120  
TCGCTCCCGG CCTTCACTCA CTCTTCCCGG CCTCGGCCC GGCTCCAG CTCTTACTT 180  
CGCGTGTCTA CAACTCAAC TCCCGGTTTC CGTGCCTCTC CACCGCTCGA GTTCTTACT 240  
CTCCATATCC GAGGGGCCCC TCCAGCATC TACCCCTCTG CCAACCTCGG GGGACCTAGC 300  
CAAGCTAGGG GGGACTGGAT CCGAGGGGTG GAGCAGCCAG GTGAGCCCGG AAAGGTGGGG 360  
CGGGGCGAGG CGGCTCCAG CCGCAGCCCG GGATCTGGTG ACGCTGGGGC TGAATTTGA 420  
CACCGAGCGG CTGCGGCGGC GGGCAGGAGG CTGCTGAGGG ATGGAGTTGG GCCCGCCCC 480  
CAGACAAGGC CCGGGGGCTC CGCCAGCAGC AGGTCCCTCG GGGCCCGAGC CTGCTGCGCA 540  
CCCGGGCTG GAGCCCAACA CCGAGGGTGG CAGACTGGCT GCCAAGGCCA CACTTTTGGC 600  
TAAAGAGGC ACTGCCAGT GTACAGTCTT GGGCATGCGC TGTTTGAGCT TCGGGGAGA 660  
GCCCAGCACT GTTCCCGGGA AAGGTGCCTA GAAGAACAAG GTGCAGGACC CCGTGTGCC 720  
TCAACAGGAG GTTGGGGGAA CAGCTCAACA ATGGCTGATG GGGCTCCTG GTGTGATAG 780  
AGATGGAAT TGGACTTGA GGCCTCTCCA CGCTGTCCA CTGCCCTGG CCTAGCGGC 840  
AGCTTGCCTT GTGGCCACCC CTGGCGCTC TGGCTCTGCT GAGCAGCTG GCAGAGGCTT 900  
CCCTGGGCTG CCGCCCGCGC AGCCCTGCCC CCGCGGAAGG CCCCCCGCTT GTCTGGGCT 960  
CCCCCGCGG CCACTGCGCG GGGGAGCGCA CGGCCGCTG GTGCACTGGA AGAGCCCGGC 1020  
GGCGCGCGCC GCAGCTTCTT GGGCGCGCG CCGCGCGCGT TGCACCCCA TCTGCTCTT 1080  
CCCGCGGGG CCGCGCGCGG CGGGCTGGGG GCGCGGCGAG CCGGCTCGG GCAGCGGGG 1140  
CGCGGGGCTG CCGCTGCGC TCGCAGCTGG TCGCGGTGG CGGCTCGGC CTGGGCCACC 1200  
GCTCGAGCA GCTGGTGGT TCCGCTTCT GCAGCGGCTC CTGCGCGCGC GCGGCTCTC 1260  
CACACGACCT CAGCCTGGCC AGCCTACTGG GCGCGGGGCG CCGCGAGCG CCGCGGGCT 1320  
CCCGCGCGT CAGCAGCCCT TCGTGGGAG CCAAGCGCTC CCAAGCGGTC TCCTTCATGG 1380  
ACGTCAACAG CACCTGGAGA ACCGTGGACC GCCTCTCGGC CACCGCTGC GGCTGCTGG 1440  
GCTGAGGGCT CGCTCCAGGG CTTTGCAGAC TGGACCCTTA CCGGTGCTC TTCCTGCTG 1500  
GGACCTTCCC GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGAGGAG GCCTCAAAGC 1560  
TGAGAGGCGC CTACCGGTGG GTGATGGATA TCATCCCGA ACAGGTGAAG GGACAACTGA 1620  
CTAGAGCGCC CAGAGCCCTC ACCCTGCGGA TCCAGCCCTA AAAGACACCA GAGACCTCAG 1680  
CTATGAGGCC CTTCGAGCCC ACTTCTACA GACTCTGGCA CTGGCCAGGC CTGAAACCTG 1740  
GGACCCCTCC TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCGCTGT 1800  
AGGGACAGCA TTTGAAGGAC ACATATTGCA GTTGCTTGGT TGAAGTGCC TGTGCTGAA 1860  
CTGGCTGTGA CTCACTATG GGAGCTGGCC CC

Seq ID NO: 605 Protein sequence  
Protein Accession #: NP\_003967.1

1 11 21 31 41 51  
MELGLGLST LSHCPWPRRQ PALWPTLAAL ALLSSVAEAS LGSAPRSPAP REGPPPVLAS 60  
PAGHLPGGRT ARWCSGRARR PPPQPSRPAP PPPAPPSALP RGGRAARAGG PGSRARAAGA 120  
RGCLRLSQLV PVRALGLGHR SDELVRFRFC SGSCRRARSP HDLSLASLLG AGALRPPPGS 180  
RPVSQPCCRP TRYEAVSFMD VNSTWRTVDR LSATACGCLG

Seq ID NO: 606 DNA sequence  
Nucleic Acid Accession #: NM\_057160.1

Coding sequence: 1..714

```
1      11      21      31      41      51
5      |      |      |      |      |      |
      ATGCCCGGCC TGATCTCAGC CCGAGGACAG CCCCTCCTTG AGGTCCTTCC TCCCAAGACC 60
      CACCTGGGTG CCTCTTTTCT CCTTGAGGCT CCACTTGGTC TCTCCGCGCA GCCTGCCCTG 120
      TGGCCCAACC TGGCCGCTCT GGCCTCTGCT AGCAGCGTGG CAGAGGCTCC CTTGGGCTCC 180
      GCGCCCGGCA GCGCTGCCCC CCGCGAAGGC CCCCCTGCTG TCTTGGGCTC CCGCGCGGCG 240
      CACCTGGCGG GGGGACGCAC GCGCGCTGG TGCAGTGGAA GAGCCCGGCG GCGCGCGGCG 300
10     |      |      |      |      |      |
      CAGCCTTCTC GGGCCGCGCC CCGCGCGCTC GCACCCCATC CTGCTCTTCC CCGCGGCGGC 360
      CCGCGCGGCG GGGCTGGGGG CCGCGGCGGC CGCGCTCGGG CAGCGGCGGC GCGGGGCTGC 420
      CGCCTGCGCT CGCAGCTGGT GCGGCTGGGC GCGCTCGGCC TGGGCCACCG CTCGACGAG 480
      CTGGTGGGTT TCGGCTTCTG CAGCGGCTCC TCGCGCGGCG CGCGCTCTCC ACACGACCTC 540
      AGCCTGGCCA GCCTACTGGG CCGCGGCGGC CTGCGACCGC CCGCGGCTC CCGCGCGGCTC 600
15     |      |      |      |      |      |
      AGCCAGCCCT GCTGCCGACC CAGCGCTAC GAAGCGTCT CCTTCATGGA CGTCAACAGC 660
      ACCTGAGGAA CCGTGGACCG CCTCTCGGCC ACCGCTGGG GCTGCTGGG CTGAGGGCTC 720
      GCTCCAGGCG TTTGAGACT GAGCCCTTAC CGTGGCTCT TCTGCTGCG GACCTCCCG 780
      CAGAGTCCCA CTAGCCAGCG GCCTCAGCCA GGGACGAAG CCTCAAGCT GAGAGGCCCC 840
      TACCGTGGG TGATGGATAT CATCCCGGAA CAGGTGAAG GACAACTGAC TAGCAGCCCC 900
20     |      |      |      |      |      |
      AGAGCCCTCA CCTGCGGATC CCGAGCTAA AAGACACAG AGACCTCAGC TATGGAGCCC 960
      TTGCGACCCA CTCTCTCAGC ACTCTGGCAC TGGCAGGCC TCGAACCTGG GACCCCTCCT 1020
      CTGATGAACA CTACAGTGGC TGAGGCATCA GCGCCCGGCC AGGCCCTGTA GGGACAGCAT 1080
      TTGAAGGACA CATATTGAGC TTGCTTGGTT GAAAGTGCCT GTGCTGGAAC TGGCCTGTAC 1140
      TCACTATGG GAGCTGGCCC C
```

Seq ID NO: 607 Protein sequence  
Protein Accession #: NP\_476501.1

```
1      11      21      31      41      51
30     |      |      |      |      |      |
      MPGLISARQG PLLEVLPPQA HLGALFLPEA PLGLSAQPAL WPTLAALALL SSVAEASLGS 60
      APRSPAPREG PPPVLASPAQ HLPGRRTARW CSRRARRPPP QPSRPAPPPP APPSALPRGG 120
      RAARAGGPGS RARAAGARGC RLRSQLVFVR ALGLGHRSD LVRFRPCSGS CRRARSPEDL 180
      SLASLLGAGA LRPPPGSRPV SQPCCRPTRY EAVSFMDVNS TWRTVDRLSA TACGCLG
```

Seq ID NO: 608 DNA sequence  
Nucleic Acid Accession #: NM\_057090.1  
Coding sequence: 29..715

```
1      11      21      31      41      51
40     |      |      |      |      |      |
      CTGATGGGCG CTCTGGTGTG TGATAGAGAT GGAAGTGGGA CTGGAGGCC TCTCCAGCT 60
      GTCCCACTGC CCTTGGCCTA GCGCGCAGGC TCCACTTGGT CTCTCCGCGC AGCCTGCCCT 120
      GTGGCCCAAC CTGGCCGCTC TGGCTCTGCT GAGCAGCGTC GCAGAGGCTC CCTTGGGCTC 180
      CCGCGCCGCG AGCCTTGGCC CCGCGCAAGG CCCCCTGCTG GTCTTGGGCT CCGCGCGGCG 240
      CCACCTGCGG GGGGACGCA CCGCCGCTG GTGCACTGGA AGAGCCCGGC GCGCGCGGCC 300
      GCAGCCTTCT CCGCGCGCGC CCGCGCGGCG TGCACCCCA 1CTGCTCTTC CCGCGCGGCG 360
      CCGCGCGGCG CCGGCTGGGG GCGCGGCGAG CCGCGCTCGG GCAGCGGGGG CCGCGGGCTG 420
      CGGCTTGGCG TCGCAGCTGG TCGCGGTGGG CGGCTCGGCG CTGGGCGACC GCTCCGACGA 480
      GCTGGTGGGT TTGCGCTTCT GCAGCGGCTC CTGCGCGGCG GCGCGCTCTC CACAGCACTC 540
      CAGCCTGGCC AGCCTACTGG GCGCGGGGCG CCGCGGAGCG CCGCGGGCTC CCGCGCGGCT 600
      CAGCCAGGCC TGCTGCGGAC CCAAGCGCTA CCAAGCGGTC TCCTTCATGG ACCTCAACAG 660
      CACCTGAGGA ACCGTGGACC GCCTCTCGCG CACCGCTGCG GGCTGCGTGG GCTGAGGGCT 720
      CGCTCCAGGG CTTTGCAGAC TGGACCCCTA CCGGTGGCTC TTCTGCTGCT GGAACCTCCC 780
      GCAGAGTCCC ACTAGCCAGC GGCCTCAGCC AGGGACGAAG GCCTCAAGC TGAGAGGCC 840
      CTACCGGTGG GTGATGGATA TCATCCCGGA ACAGGTGAAG GGACAACTGA CTAGCAGCCC 900
      CAGAGCCCTC ACCCTGCGGA TCCAGGCTTA AAGACACCA GAGACCTCAG CTATGGAGCC 960
      CTTGGAACCC ACTTCTCACA GACTCTGGCA CTGGCCAGGC CTGAACTCTG GAGCCCTCCT 1020
      TCTGATGAAC ACTACAGTGG CTGAGGCATC AGCCCGCGCC CAGGCCCTGT AGGGACAGCA 1080
      TTTGAAGGAC ACATATTGCA GTTGTCTGGT TGAAGTGCCT TGTGCTGGA 1140
      CTCACATGAG GAGCTGGGCC C
```

Seq ID NO: 609 Protein sequence  
Protein Accession #: NP\_476431.1

```
1      11      21      31      41      51
65     |      |      |      |      |      |
      MELGLGGLST LSHCPWPRRQ APLGLSAQPA LWPTLAALAL LSSVAEASLG SAPRSPAPRE 60
      GPPPVLASPA GHLPGGRTAR WCSGRARRPP PQPSRPAPPP PAPPALPRG GRAARAGGPG 120
      SRARAAGARG RLRSQLVFVR RALGLGHRSD ELVRFRPCSG SCRRARSPHD LSLASLLGAG 180
      ALRPPPGSRP VSQPCRPTRY YEAVSFMDVNS STWRTVDRLS ATACGCLG
```

Seq ID NO: 610 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1746

```
1      11      21      31      41      51
80     |      |      |      |      |      |
      ATGCCACTGA AGCATTATCT CCTTTTGTCT GTGGGCTGCC AAGCCTGGGG TGCAGGGTTG 60
      GCTTACCATG GCTGCCCTAG CGAGTGTACC TGCTCCAGGG CCTCCAGGT GGAGTGCAAC 120
      GGGGACGACA TTGTGGCGGT GCCCACCCTC CTGCCCTGGA ACGCCATGAG CCTGCAGATC 180
      CTCACACGCG ACATCACTGA ACTCAATGAG TCCCGGTTCC TCAATATCTC AGCCCTCATC 240
      GCCTTGAGGA TTGAGAAGAA TGAGCTGTGG CGCATCACGC CTGGGCGCTT CCGAAACCTG 300
      GGCTCGCTGC GCTATCTCAG CCTCGCCAAC AACAGCTGCG AGGTTCTGCC CATCGGCTTC 360
85     |      |      |      |      |      |
      TTCCAGGGCC TGGACAGCCT TGAGTCTCTC CTCTGTGCCA GTAAACGACT GTTGACAGATC 420
      CAGCGGCCCC ACTTCTCCCA GTGCAGCAAC CTCAGGAGCG TGCACTGGA CCGCAACCAT 480
      CTGGAATACA TCCCTGACGG AGCCTTCGAC CACCTGGTAG GACTCACGAA GCTCAATCTG 540
```

	GGCAAGAATA	GCCTCACCCA	CATCTCACCC	AGGGTCTTCC	AGCACCTGGG	CAATCTCCAG	600.
	GTCTCCGGC	TGTATAGAGAA	CAGGCTCACG	GATATCCCCA	TGGGCACITT	TGATGGGCTT	660
	GTAACTGCG	AGGAACCTGG	TCTACAGCAG	AACCAGATTG	GACTGCTCTC	CCCTGGTCTC	720
5	TTCCACAACA	ACCACAACCT	CCAGAGACTC	TACCTGTCCA	ACAACCACAT	CTCCAGCTG	780
	CCACCCAGCA	TCTTCATGCA	GCTGCCCCAG	CTCAACCCGT	TTACTCTCTT	TGGGAATTC	840
	CTGAAGGAGC	TCTCTCTGGG	GATCTTCGGG	CCCATGCCCA	ACCTGCGGGA	GCTTTGGCTC	900
	TATGACAACC	ACATCTCTTC	TCTACCCGAC	AATGTCTTCA	GCAACCTCCG	CCAGTTGCAG	960
	GTCTGTATTC	TTAGCCGCAA	TCAGATCAGC	TTCTCTCC	CGGGTGCTT	CAACGGGCTA	1020
10	ACGGAGCTTC	GGGAGCTGTG	CCTCCACACC	AACGCACCTG	AGGACCTGGA	CGGGAATGTC	1080
	TTCCGCTATG	TGGCCAACTC	GCAGAACATC	TCCCTGCAGA	ACAATCGCCT	CAGACAGCTC	1140
	CCAGGGAATA	TCTTCGCCAA	CGTCAATGGC	CTCATGGCCA	TCCAGCTGCA	GAACCAACCAG	1200
	CTGGAGAACT	TGCCCTCGGG	CATCTTCGAT	CACCTGGGGA	AACTGTGTGA	GCTGCGGCTG	1260
	TATGACAATC	CCTGGAGGTG	TGACTCAGAC	ATCCTTCGCG	TCCGCAACTG	GCTCCTGCTC	1320
15	AACCAGCCTA	GGTTAGGGAC	GGACACTGTA	CCTGTGTGTT	TCAGCCGAGC	CAATGTCCGA	1380
	GGCCAGTCCC	TCATTATCAT	CAATGTCAAC	GTGTGCTGTC	CAAGCGTCCA	TGTCCCTGAG	1440
	GTGCTAGATT	ACCCAGAAAC	ACCATGGTAC	CCAGACACAC	CCAGTTACCC	TGACACCACA	1500
	TCCGTCTCTT	CTACCACTGA	GCTAACCCAG	CCTGTGGAAAG	ACTACACTGA	TCTGACTACC	1560
	ATTCAAGTCA	CTGATGACCG	CAGCGTTTGG	GGCATGACCC	AGGCCAGAG	CGGGCTGGCC	1620
20	ATTGCGGCCA	TGTAAATTGG	CATTGTGGCC	CTGGCCTGCT	CCCTGGCTGC	CTGCGTCGCG	1680
	TGTGTGCTGT	GCAAGAGAG	GAGCCAGCT	GTCTGTATGC	AGATGAAGGC	ACCCAAATGAG	1740
	TGTTAAAGAG	GCAGGCTGGA	GCAGGCTGG	GGAAATGATG	GACTGGAGGA	CCTGGGAATT	1800
	TCATCTTCTT	GCCTCCACCC	CTGGTCCAT	GGAGCTTCC	CGTATTGCT	CTTCTGCGCC	1860
	CTAGATAAAG	GTGTGCTTAC	CTCTCTCTGA	CTTGCCTGAT	TCTCCGCTAG	AGAAGCAGGT	1920
25	CGTGCAGGAC	CTTCTACAA	TCAGGAAGAT	AGATCCAACT	GGCCATGGCA	AAAGCCCTGG	1980
	GGATTTCGGA	TTCATACCCC	TGGGCTTCTT	TCGAGAGGGC	TCTTCTCCA	AATCCTCCCC	2040
	ACCTGTCTCT	CAAGAACAGC	CTTCCCTGCG	CCCAGGCCCC	CTCGGGCTCT	CTGTAGACTC	2100
	AGTTAGTCCA	CAGCCTGTCT	ACTTGTGGGG	AATAGTTCTC	CGCTGAGATA	GCCCCCTCG	2160
	CCTAAGTATT	ATGTAAGTTG	ATTTCCCTTC	TTTTGTTTCT	CTTGTGTTG	CTATGGCTTG	2220
30	ACCCAGCATG	TCCCTCAAA	TGAAAGTTCT	CCCCTTGATT	TTCTGCTCCT	GAAGGCAGGG	2280
	TGAGTTCTCT	CCTCAAGAAA	GACTTCAAA	CATTTAACCTG	GTTCCTTAAG	AGCCGTCAAT	2340
	CAGCCTGGTT	TGCGGATGTC	TATGAAAGAG	AGAAGGAAAA	TCATGCCGCT	CAGTTCTCTG	2400
	AGACAGAGAG	GCGGTCTATC	GTGTCTCACT	TGTGATTTTT	ATCTGGAAAA	GGAGAAAAA	2460
	CCCCAGCACA	GCAAGCTCAG	CCTTTAGAG	AAGGATATTT	CCAACTGCA	AACCTTGCTT	2520
35	TGAAAGTTT	AGCCCTTTAA	GSAAATGAA	CATGTAGAA	TTTGGAATTC	TAAAAACATT	2580
	AAATCAGCT	TATTAAATAG	GGATAGAGAA	AGAAATCTGG	TGCGTGGGG	TCCCTGTGTT	2640
	CACCCCTAGA	GTTTGTGTTA	AAATTTTAA	TTGAAGCATG	TGAAGTGTAC	STGCAGAAAA	2700
	GTGGGAAAT	GATAGTGTAT	GGCTTGGTGG	ATTTTCAACA	ACTGAACATA	CCTGTGTAA	2760
	CAGCATCTAG	ACCCAGACCC	AGAGCATCAC	AAATATCCCC	CATCCTGGGC	TTTTCCAGAA	2820
40	GGAGATGGGG	GCTTCTGAAG	ATGGACTTAC	CTGGGACCTG	CCCCCATGA	GCCAGGACGG	2880
	TCCCCCCACA	GTGAGCCTGT	GCAAGGCCCC	CGTGGCCAGG	GGTGGAGGAG	AATATGTGG	2940
	TGTGACAGGG	ATGGGAGACT	GTGGCCTGAA	CAGGAGATT	TATTATATCT	GGAGACCTG	3000
	AGAGACCTTG	AGACCTGGGG	CACCATGGCT	GGCCAGGTCA	GAAGCATCCT	GACTGCAGAG	3060
	GTCCGTGACG	CCAACCTCTC	TTCCCTGGCA	GCAAGTTGTC	TGCGGCTCAT	CGGAGGCCCC	3120
45	TCCGCTCGGA	GCTTCTATG	GAOGTATAT	GCCTGTATCT	GTTTTAAAT	TTCTTCTTC	3180
	ACTTAGGGGA	AGTGAATCG	CTCAGAGATG	AGATCCTTTA	ATTGAAAAAG	AAGTGAACG	3240
	GAATCTAGTG	TCTTTCTAAT	GTGGTAAAT	TCTCCATCAA	CATCACAGTC	AGCTGGCAGC	3300
	TGAACCTCAG	AACTCTCACT	ACAGCAGGCG	ACACGGGGGT	ACACCGATGG	GTCACTATGG	3360
	GTCTGGGGGC	TCCCTCTGCG	TGTGGTCTGG	TTAGGAGTTG	AGTTGTTTGC		3420
50	TCCAGGGTTA	TTCTCTCTCT	CGAGTCACAG	TCACAGCAAT	ACCTGCCTTC	TCTGGCTTTC	3480
	CTGCTATACA	CATATTCACA	TGGCGCTCAA	GAAGTTAGGC	TCATGGCAAC	GTGTGTCTTT	3540
	CTCTGGACAA	CTGGCCAGT	TTACAGTGAA	ATGAGAAAT	TCAGGTCTCC	ACGTCTGCCC	3600
	AGGAAAGAAC	TTCACTGAC	TCCACGGGGA	TCTGGAATTC	CAAGCCCAAT	CCCGATCGGC	3660
	TCTTATTAGC	TCCCGCTGCC	ACAAGACACC	TGTGCTTTGG	AAATCCACCA	CCAATCCCGA	3720
55	TGCGCTCTTA	TTAGCTCCCC	GCTCCACAAG	ACACCTGTGA	TCTGGAAATC	TACCACCAAT	3780
	CCGATTCGGC	TCTTATTAGC	TCCCGCTCC	ACAAGACACC	TGTGACATCC	TCCAGGGCCA	3840
	CAGGAGCAG	TGCTGACCAG	TTTTCCCTTC	CAGTTCTGTC	ACAAAAGTG	TCCAGAGGGC	3900
	TGTTTGCAAA	CATAGTGCA	CTTTGTAGCT	TTTCAACCTC	TGTCCAGGG	AATCTAGGAG	3960
	AGATAGAGCC	CGCTAGAGTC	AAGAGATGTC	ATCCCCCAG	GGTCTCCAAG	GCATTTCCAC	4020
60	ACTATTGGTG	GCACTTGAGC	GACATGCACC	AAGGCTTGCC	AGAGCCAAAC	GGAAGTGAGC	4080
	CCAGAGCATG	GCACATGAGC	ATCACCGCT	GATGGTGGCC	TGCTGTGCTC	GGTGCCAAAC	4140
	GGGGCATCCC	GGCCGTTACC	CCTCCAGACA	GGAAGCATGG	GTTCGCCAC	AGACCTGTG	4200
	GGTGTCTCTG	TGAGTGGGCT	CCAGATGTCT	TTGTGCATAG	GCACAAGTGG	GCCAGGGCTG	4260
	GAGGGAGGTG	GGAAACCTCA	TCATCCGGTG	GGCCCTGCCA	ATCTTAACCC	AGAACCTTTA	4320
65	GGTATTCTCTG	GCACTAGCCA	TGACATTGGA	GCACCTTCTC	CTCCAGCCAG	AGGCTGACCT	4380
	GAGGGCCACT	GTCTCAGAT	GACACCAACC	AGGAGCACCC	TAGGTGAGGG	GTGAGGGCCC	4440
	CCTTATGTGA	ACCTCTTGCC	TCTTCTTTTC	TCCATCAGA	GTGGTTGGAT	GGAGCCATTG	4500
	GCCTCTCTTT	CTTCAGCGGG	CCCTTCAACC	TCTTGCACC	ATGTTGTCTG	CTGAGGAGC	4560
	TACTAGAAAA	GCTGAGTGGG	GTCTCTTTTC	CAACAGGATG	ATGCATTTGC	TCAATTCTCA	4620
70	GGGCTGGAAT	GAGCCGGCTG	GTCCCCAGA	AAGCTGGAGT	GGGTACAGA	GTTCAGTTT	4680
	CCTCTCTGTT	TACAGCTCTT	TGACAGTCCC	ACGCCATCT	GGAGTGGGAG	CTGGGAGTTA	4740
	GTGTTGGAGA	AGAAACAACA	AAAGCCAAAT	AGAACCATA	TTTTTAAAA	GTGCTTACTG	4800
	TGCACAGATA	CTCTTCAAGC	ACTGGAAGTG	GATTCTCTCT	CTAGCCCTCA	GCACCCCTGC	4860
	GGTAGGAGTG	CCGCTCTTAC	CCACTTGTGA	TGGGGTACAG	AGGCACCTGC	TCTTCTGCAT	4920
75	GGTGTTCAT	AGGCTGGGAG	TTTTATTAT	CTCTTCAAAC	TTTGTACAAG	AGCTCATGGC	4980
	TTGTCTGGG	CTTCTGTCAT	TAAACCAAG	GAAATGGAAG	CCATTCCCTC	GTGCTCTCC	5040
	TTAGTCTTGG	TCATCAGAAC	CTCACTTGGT	ACCATATAGA	TCAAAGCCTT	TGTAACCA	5100
	GGAAAAATA	AACTCTTCCA	TCCCTTAAAG	AATAGAATAG	TTTGTCCCTC	TCATGGGAAT	5160
	TGGCTGTAT	GTATATTGTT	CTTCTCTCTT	AGAAATTAGA	GATACAAGAG	TTTACTTAG	5220
80	AACTTTTCAT	GGACCAAT	TCCACAACCT	TTCAGATGCT	GATGTAGAGC	TATTGGGAAA	5280
	GAACCTTCAA	ACTCAGGAG	TTTGACAGAG	GCAGACAGCT	AGAGATAACT	CGGAGCCAG	5340
	AGTTGTGGA	CAGATGTAG	ATGATCTCTA	GCTTTTAGCC	ATAAACCACT	CAAAGATTCA	5400
	GCCCCAGAT	CCACAGTCA	GAACGAATC	TGCGTTGTTG	GGAAGCCAGC	AGTGGCTTG	5460
	GGAAAGGAGC	CATGGCTGTG	GTTCAGAGAG	GGTGGGCTGG	CAAGCCACTT	CCGGGGAATA	5520
85	CTCTTCCGCG	CCAGGTTTC	TTCTTCTCTT	AAGGAGAGAT	TGTTCTCACC	AACCCGCTGC	5580
	CTTCATGCTG	GCTTCAAAGC	TAGATCATGT	TTGCTTGTCT	TAGAGAATTA	CTGCAATCA	5640
	GCCCCAGTGC	TGGCGATGCG	ATTTACAGAT	TTCTAGGCC	TCAGGGTTT	GTAGAGTGTG	5700
	AGCCCTGGTG	GGCAGGTTG	GGGGTCTG	CTTCTGCTGG	ATGCTGCTTG	TAATCCATT	5760

GGTGACAGA ATCAACAATA AATAATATAC ATGTAT

Seq ID NO: 611 Protein sequence  
Protein Accession #: BAB84587.1

1 11 21 31 41 51  
MPLKHYLLLL VGCQAWGAGL AYHGCPSECT CSRASQVECT GARIVAVPTP LPWNAMSLQI 60  
LNTHITELNE SPFLNISALI ALRIEKNELS RITPGAFRNL GSLRYLSLAN NKLQVLPGL 120  
10 FQGLDSLESLL LLSNQLLQI QPAHFSQCSN LKELQLHGNH LEYIPDGAFL HVLGLTKLNL 180  
GKNSLTHISP RVFQHLGNLQ VLRLYENRLT DIPMGTFDGL VNLQELALQQ NQIGLLSPGL 240  
PHNNHNLRL YLSNNHISQL PPSIFMQLPQ LNRLTLFGNS LKELSLGIFG PMPNLRELWL 300  
YDNHISLPLD NVFSNLRQLQ VLILSRNQIS FISPGAFNGL TELRELSLHT NALQDLGDNV 360  
15 FRMLANLQNI SLQNNRLRQL PQNIPANVNG LMAIQLQNNQ LENLPLGIFD HLGKLCLELRL 420  
YDNPWRCDSL ILPLRNWLLL NQPRLGTDIV PVCFSANVR GQSLIINVN VAVPSVHVPE 480  
VPSYPETPWY PDTSPYDPTT SVSSTELTTS PVEDYDLTT IQVTDDRVSUW GNTQAQSGLA 540  
IAAIVIGIVA LACSLAACVG CCCCCKRSQA VLMQMKAFNE C

Seq ID NO: 612 DNA sequence  
Nucleic Acid Accession #: XM\_098151  
Coding sequence: 1..447

1 11 21 31 41 51  
ATGATGCATT TGCTCAATTC TCAGGGCTGG AATGAGCCGG CTGGTCCCC AGAAAGCTGG 60  
AGTGGGTAC AGAGTTCAGT TTCTCTCTCT GTTTACAGCT CCTTGACAGT CCCACGCCCA 120  
25 TCTGGAGTGG GAGCTGGGAG TCAGTGTGGG AGAAGAAACA ACAAAGCCA ATTAGAACA 180  
CTATTTTAA AAGTGTCTTA CTGTGCACAG ATACTCTTCA AGCACTGGAC GTGGATTCTC 240  
TCTCTAGCCC TCAGACCCCC TGCGGTAGGA GTGCCGCCCTC TACCCACTTG TGATGGGSTA 300  
30 CAGAGGCACT TGCTCTTCTG CATGGTGTTC AATAGGCTGG GAGTTTATT TATCTCTTCA 360  
AACTTTGTAC AAGAGCTCAT GCCTTGTCTT GGGCTTTCTG CATTAAACA AAGGAAATGG 420  
AAGCCATTCC CCTGTGCTG TCCTTAG

Seq ID NO: 613 Protein sequence  
Protein Accession #: XP\_098151

1 11 21 31 41 51  
MMHLLNSQGW NEPAGPPESW SGVQSSVFLS VYSSLTVPRP SGVGAGSQCW RRNKSQLEP 60  
40 LFLKSAVCAQ ILFKHWIWL SLALSTPAVG VPPLPTCDGV QRHLLFCMVF NRLGVLFISS 120  
NPVQELMACL GLSSLNQRKW KFPFCCSP

Seq ID NO: 614 DNA sequence  
Nucleic Acid Accession #: NM\_002658.1  
Coding sequence: 77..1372

1 11 21 31 41 51  
GTCCCCGCGC GCGCTCGCG CCTCTCTGCT GCAGGCCACC GAGGCCGCCG CGCTCTAGCG 60  
50 CCCCGACCTC GCCACCATGA GAGCCCTGCT GCGCGCCCTG CTCTCTGCGC TCCTGGTCTG 120  
GAGCGACTCC AAAGGACGCA ATGAATTCCA TCAAGTTCCA TCGAAGTGTG ACTGTCTAAA 180  
TGGAGGAACA TGTGTGTCCA ACAAGTACTT CTCCAACTAT CACTGTGTGA ACTGCCCAAA 240  
GAAATTGCGA GCGCAGCACT GTGAAATAGA TAAGTCAAAA ACCTGCTATG AGGGGAATGG 300  
55 TCAGTTTAC CGAGGAAAGG CCAGCACTGA CACCATGGGC GCGCCCTGCC TGCCCTGGAA 360  
CTCTGCCACT GTCTCTCAGC AAACGTACCA TGCCCAACA TCTGATGCTC TTCAGTCTGG 420  
CCTGGGAAA CATAATTACT GCAGGAACCC AGACAACCCG AGGCGACCCG GGTGCTATGT 480  
GCAGGTGGGC CTAAAGCCGC TTGTCCAAGA GTGCATGGTG CATGACTCGC CAGATGGAAA 540  
60 AAAGCCCTCC TCTCTCCACG AAGAATTAAA ATTTCACTGT GGCCAAAAGA CTCTGAGGCC 600  
CGCTTTAAG ATTATTGGGG GAGAATTCAC CACCATCGAG AACCCAGCCG GTTTGCGGCG 660  
CATCTACAGG AGGCACCGGG GGGGCTCTGT CACCTACGTG TGTGAGGCA GCCTCATCAG 720  
65 CCTTGTGG GTGATCAGCG CCACACACTG CTTCATTGAT TACCCAAAGA AGGAGGACTA 780  
CATGCTCTAC CTGGGTGCTC CAAGGCTTAA CTCCAACAG CAAGGGGAGA TGAAGTTTGA 840  
GGTGGAAAAC CTCATCTTAC ACAAGGACTA CAGCGCTGAC ACCTTGTCTC ACCACAACGA 900  
CATTGCCCTG CTGAAGATCC GTTCCAAGGA GGCAGGCTGT GCGCAGCCAT CCCGACTAT 960  
70 ACAGACCATC TGCTGCCCCG CGATGTATAA CGATCCCCAG TTTGSCACAA GCTGTGAGAT 1020  
CACTGGCTTT GGAAAAGAGA ATTCTACOGA CTATCTCTAT CCGGAGCAGC TGAATATGAC 1080  
TGTTGTGAAG CTGATTTCCT ACCGGGAGTG TCAGCAGCCC CACTACTACG GCTCTGAAGT 1140  
CACCACAAA ATGCTATGTG CTGCTGACCC CCAATGGAAA ACAGATTCTT GCCAGGGAGA 1200  
CTCAGGGGGA CCGCTGCTCT GTTCCCTCCA AGGCCGCGATG ACTTTGACTG GAATTGTGAG 1260  
75 CTGGGCGCTG GATGTGCCCC TGAAGGACAA GCCAGCGCTC TACACGAGAG TCTCACACTT 1320  
CTTACCTGG ATCCGCAGTC ACACCAAGGA AGAGAATGSC CTGGCCCTCT GAGGGTCCCC 1380  
AGGGAGGAAA CCGGCACAC CCGCTTCTT GCTGTGTGTC ATTTTTCGAG TAGAGTCATC 1440  
TCCATCAGCT GTAAGAAGAG ACTGGGAAGA TAGGCTCTGC ACAGATGGAT TTGCTGTGG 1500  
CACCACAGG GTGACGACA ATAGCTTTAC CCTCAGGAT AGGCCTGGT GCTGGCTGCC 1560  
80 CAGACCCCTC GGCCAGGATG GAGGGGTGGT CCTGACTCAA CATGTTACTG ACCAGCAACT 1620  
TGCTTTTTTC TGGAGTGAAG CCTGAGGAG TTAAGAGGG CAGGCACTCT CCTGTGCATG 1680  
GGCTCGAAGG GAGAGCCAGC TCCCCGACC GGTGGGCATT TGTGAGGCC ATGCTTGAGA 1740  
AATGAATAAT TTCCCAATTA GGAAGTGTAA GCAGCTGAGG TCTCTTGAGG GAGCTTAGCC 1800  
AATGTGGGAG CAGCGTTTG GGGAGCAGAG AACTAACA CTTGAGGCA GGGCTCTGAT 1860  
85 ATTCCATGAA TGTATCAGGA AATATATATG TGTGTGTATG TTTGCACTT TGTGTGTGG 1920  
GCTGTGAGTG TAAGTGTGAG TAAGAGCTGG TGCTGTATTG TTAAGTCTAA ATATTCTCT 1980  
AAACTGTGTG GACTGTGATG CCACACAGAG TGCTCTTCTT GGAGAGGTTA TAGGCTACTC 2040  
CTGGGCGCTC TTGGGTCCCC CAGTGACAG TGCCCTGGGA TGTACTTATT CTGACAGCAT 2100  
ACCTGTGACC AGCACTGTCT CAGTTTCACT TTCACATAGA TGTCCTTTC TTGGCCAGTT 2160  
ATCCCTTCTT TTAGCCTTAG TTCAATCAAT CCTCACTGGG TGGGGTGAAG ACCACTCCTT 2220  
ACACTGAATA TTTATATTTT ACTATTTTAA TTTATATTTT TGAATTTTAA AATAAAGTG 2280  
ATCAATAAAA TGTGATTTTT CTGA

Seq ID NO: 615 Protein sequence  
Protein Accession #: NP\_002649.1

```

5      1      11      21      31      41      51
|      |      |      |      |      |
MRALLARLLL CVLVVSDDSKG SNEHQVPSN CDCINGGTCV SNKYFSNIHW CNCPKFGGQ 60
HCEIDKSKTC YEGNGHPLYRG KASTDTMGRP CLPWNSTATVL QQTYHAHRSD ALQLGLGKHN 120
YCRNPDNRNR PWCVYQVGLK PLVQECMVHD CADGKCPSSP PEELKFPQCGQ KTLRPRFKII 180
10 GGEFTTIENQ PWFPAIYRRH RGSSTVYVCG GSLISPCWVI SATHCPIIDYP KKEDYIVYLG 240
RSRLNSNTQG ENKFEVENLI LHKDYSADTL AHENDIALLK IRSKEGRCAQ PSRTIQTICL 300
PSMYNDPQFG TSCEITGPGK ENSTDYLYPE QLKMTVVVKLI SHRECQPHY YGSEVTTKML 360
CAADPQWKTD SQQDSGGPL VCSLQGRMTL TGVISWGRGC ALKDKPGVYT RVSHFLPWIR 420
SHTKEENGLA L

```

Seq ID NO: 616 DNA sequence  
Nucleic Acid Accession #: NM\_024422.1  
Coding sequence: 202..2907

```

20      1      11      21      31      41      51
|      |      |      |      |      |
CGCCAAAGGA AAAGCCCCTT GGATGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAGCAC 60
CTCTCOGCGC GCCCCACCTC CTCCGCTCTG CGCTCCTCTC GAGCAGCGGG CCCGACTGCG 120
GCTCCGCGCG CGCCCTCTCG CCGCGGAGGC CCTCCTACCC CGGCCCGACG CTCGCCCGCG 180
25 GACCTGCCCC GAGCCTCTC CATGAGAGCA GCCCGCCCTC CCGCTCTCTG GAACGGAGCC 240
CTCTCGCGCG TGCTCTCTG GACCTCGCG ATCTTAATAT TTGCCAGTGA TGCCCTGCAA 300
AATGTGACAT TACATGTGCC CTCCAACTA GATGCGGAGA AACTTGTTGG TAGAGTTAAC 360
CTGAAAGAGT GCTTTACAGC TGCAAACTA ATTCAATCAA GTGATCCTGA CTTCCAAATT 420
TTGGAGGATG GTTCAGTCTA TACAACAAAT ACTATTCTAT TGTCTCTCGA GAAGAGAAGT 480
30 TTTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATATT TGTCTTTTGG 540
GAGCATCAAA CAAAGTCTCT AAAGAAAAGA CATACTAAAG AAAAAGTTCT AAGCGCGCGC 600
AAGAGAAGAT GGGCTCCAAT TCCTTGTTGG ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA 660
CTTTTCTTTC AACAGTTCTA ATCTGACACG GCCCAAAACT ATACCATATA CTATTCCATA 720
AGAGGTCTCG GAGTTGACCA AGAACTCGG AATTATTTTT ATGTGGAGAG AGACACTGGA 780
35 AACTTGTATT GTACTCGTCC TGTAGATCGT GAGCAGTATG AATCTTTTGA GATAATTGCC 840
TTTGCAACAA CTCCAGATGG GTATACTCCA GAACCTCCAC TGCCCTTAAT AATCAAAATA 900
GAGGATGAAA ATGATACTA CCAATTTTTT ACAGAAAGAA CTTATACTTT TACAATTTTT 960
GAAAATTGCA GAGTGGGCAC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1020
GACACGATGC ACACACGCGT GAAGTACTCC ATCATTGGGC AGTGGCCACC ATCAGCCACC 1080
40 CTATTTCCTA TGCACTCAAC TACAGGCGTG ATCACCACAA CATCATCTCA GCTAGACAGA 1140
GAGTTAATTG CAAAGTACCA GTTGAAAATA AAAGTACAAG ACATGGATGG TCAGTATTTT 1200
GGTCTACAGA CAACTTCAAC TTGTATCATT AACATTGATG ATGTAATGTA CCACCTGCCA 1260
ACATTACTCT GTACTTCTTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGGAATC 1320
45 TTACGAGTTA CTGTTGAGGA TAAGGACTTA GTGAATCTG CTAAGTGGAG AGCTAATAT 1380
ACCATTTTAA AGGGCAATGA AAATGGCAAT TTTAAATTTG TAACAGATGC CAAAACCAAT 1440
GAAGGAGTTC TTTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG 1500
CAAAATGGTG TAGTTAATGA AGCTCCATTT TCCAGAGAGG CTAGTCCAAG ATCAGCCATG 1560
AGCAGACGAA CAGTTACTGT TAATGTAGAA GATCAGGATG AGGGCCCTGA GTGTAAACCT 1620
50 CCAATACAGA CTGTTCCGAT GAAAGAAAAT GCAGAAGTGG GAACAACAAG CAATGGATAT 1680
AAAGCATATG ACCCAGAAAC AAGAGTAGC AGTGGCATAA GGTATAAGAA ATTAAGTAT 1740
CCAAAGGGT GGGTCACCAT TGATGAAAAT ACAGGATCAA TCAAGTTTTT CAGAAGCCTG 1800
GATAGAGAGG CAGAGACCAT CAAAATGGC ATATATAATA TTACAGTCTT TGCATCAGAC 1860
CAAGGAGGGA GAACATGTAC GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1920
55 AGCCCATTC AACTTAAAAA GACAGTATC ATCTGCAAA CCAACCATGC ATCTGCGGAG 1980
ATTGTTGGCG TTGATCTCTA TGAGCCTATC CATGGCCAC CCTTTGACTT TAGTCTGGAG 2040
AGTTCTACTT CAGAAGTACA GAGAATGTGG AGACTGAAAG CAATTATGTA TACAGCAGCA 2100
CGTCTTTCTT ATCAGAAATG TCCTCCATTT GGCTCATATG TAGTACCTAT AACAGTGAGA 2160
GATAGACTTG CAGTGTCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCATTACC 2220
60 GAAAATGACT GCACACATCG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGGA 2280
AAGTGGGCCA TCCTTGCAAT ATTGTTGGGC ATAGCAATGC TCTTTTGCA CTGTTTACG 2340
CTGGTCTGTG GGGCTTCTGG GACGCTTAAA CAAACCAAG TAATTCCTGA TGATTAGCC 2400
CAGCAGAAC TAATTGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGCG 2460
AATGGCTTCA CAACCCAAAC TGTGGCGCTG TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2520
65 TCAGGAATCA AAAACGAGG TCAGGAGACC ATCGAAATGG TGAAAGGAGG ACACAGAGCC 2580
TCGGAATCCT GCCGGGGGCG TGGCCACCAT CACACCTGG ACTCCTGCAG GGGAGGACAC 2640
ACGGAGGTGG ACAACTGCGC ATACACTTAC TCGGAGTGGC ACAGTTTAC TCAGCCCGT 2700
CTTGGTGAAG AAGTGTATCT GTGTAAATCA GATGAAAATC ACAAGCATGC CCAAGCATAT 2760
70 GTCTTGACAT ATAACATGTA AGGAAGAGGA TCGGTGGCTG GGTCTGTAGG TTGTTGAGT 2820
GAACGACAAG AAGAAGATGG GCTTGAATTT TTGATAATT TGGAGCCCAA ATTTAGGACA 2880
CTAGCAGAAG CATGCATGAA GAGATGAGTG TGTCTAATA AGTCTCTGAA AGCCAGTGGC 2940
TTATGACTT TTAACAAAAA TTACAAACCA AGAATTTTTT AAAGCAGAAG ATGCTATTG 3000
TGGGGTTTTT TCTCTCATT TTTGGATGGA ATCTCTTGG TCAATGTCAC ATTTACAGAG 3060
75 AGACACTATA AACAAATGTA CAAATTTTTT AATTTTACA TATTTTAAA TTACTTATCT 3120
TCTATCCAG GAGGTCTACA GAGAAATTA AGTCTGCGCT ATTTGTTACA TTTGGGTATA 3180
ATGACACAG CCAATTTATA GTGCAATAAA ATGTAATTA TTCAAGTCTT TATTATAGAC 3240
TATTTGAGC ACAACCTAAT GGAATTTGT AGAGACCTTG CTTTAACATT ATCTCCAGT 3300
AATTAAGTGT TCATGTGGTG CTGGAAGACT GTTGTTTTCC TGAACATCTA AAGTCTGTAG 3360
ACTGCATTCT TGCTATTATT TTATTCTTGT AATGTGACCT TTTCCTGTG CAAAGGGAGA 3420
TTCTAGCCA GGCATTGACT ATTACAATTT CATT

```

Seq ID NO: 617 Protein sequence  
Protein Accession #: NP\_077740.1

```

85      1      11      21      31      41      51
|      |      |      |      |      |
MEARPSGSW NGALCRLLLL TLAILIFASD ACKNVTLHVP SKLDAEKLVG RVNLKECFDA 60
ANLIHSSDPD FQILEDGSVY TTNITLLSSE KRSPTILLSN TENQEKKIF VFLHEQTKVL 120

```



WO 02/086443

KKRHTKEKVL RRAKRRWAPI PCSMLENSLG PFPLFLQVQV SDTAQNYTIY YSIRGPGVDQ 180  
 EPRNLFYVER DTGNLYCTRP VDRQYVESFE IIAFATTPDG YTPLEPLPLI IKIEDENDNY 240  
 PIFTEETVTF TIFENCVRGT TVGQVCATDK DEPTMTMTRL KYSIIGQVFP SPTLFSEHPT 300  
 TGVITTSQ LDRLEIDKQV LKIKVQMDG QYFGLQTTST CIINIDVMD HLPFTFRTSY 360  
 VTSVEENTVD VEILEVTVED KDLVNTANWR ANYTILKGNE NGNFKIVTDA KYNIEGVLGV 420  
 KPLNYEEKQQ MLQIGVVNE APFSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVRM 480  
 KENAEVGTTS NGYKAYDPET RSSSGIRYKK LTDFTGNVTI DENTGSIKVF RSLDREARTI 540  
 KNGIYNITVL ASDQGGRTCT GTLGIILQDV NDNSFFIPKK TVIICKPTMS SAEIVAVDPD 600  
 EPIHGPPFDF SLESSTSEVQ RMWRLKAIND TAARLSYQND PPFQSVVVPV TVRDRLGMS 660  
 VTSLDVTLCD CITENDCTHR VDPRIQGGGV QLGKWAIALI LLGIALLPFI LFTLVCGAGS 720  
 TSQPKVPIPD DLAQONLIVS NTEAPGDDKV YSANGFTTQT VGASAGQVCG TVGSGIKNGG 780  
 QETIEMVKGQ HQTSSECRGA GHHHTLDSR GGHTEVDNCR YTYSEHSTP QPRLGEKVYL 840  
 CNQDENHKA QDYVLTYNIE GRGSVAGSVG CCSERQEDG LEFLDNLEPK FRTLAECMK 900  
 R

Seq ID NO: 618 DNA sequence  
 Nucleic Acid Accession #: NM\_004949.1  
 Coding sequence: 202..2745

1 11 21 31 41 51  
 CGCCAAAGGA AAAGCCCTT GGATGAGAGG CAGGCGCTTC AGAGAAGCTA AGAAAAGCAC 60  
 CTCTCCGGCG GCGCCACCTC CTCCGCTCG CGCTCCTCCT GAGCAGCGGG CCCGAGCTGC 120  
 GCTCCGGCGG CGGCGCTCGC CCGCGGAGC CCTCTACCC CGGCGCGAGC CTGCGCCGCG 180  
 GACCTGCCCG GAGCCCTCTC CATGGAGGCA GCGCGCCCTC CGGCTCCTG GAACGGAGCC 240  
 CTCTGCGCGG TGCTCTCTCT GACCTCTCGG ATCTTAATAT TTGCCAGTGA TGCCCTGCAA 300  
 AATGTGACAT TACATGTTCC CTCCAACTA GATGCCGAGA AACTTGTGG TAGAGTTAAC 360  
 CTGAAGAGT GCTTTACAGC TGCAATCTA ATTCATTCAA GTGATCTCGA CTCCAAAT 420  
 TTGGAGGATG GTTCACTCTA TACAACAAT ACTATTCTAT TGCTCTCGGA GAAGAGAAGT 480  
 TTACCATAT TACTTTCCAA CACTGAGAAC CAAGAAAAGA AGAAAATATT TGCTTTTGTG 540  
 GAGCATCAAA CAAGGTCTCT AAAGAAAAGA CATACTAAG AAAAAGTTCT AAGCGCGGCC 600  
 AAGAGAAGAT GGGCTCCAAAT TCCTTGTTCG ATGCTAGAAA ACTCCTTGGG TCCTTTTCCA 660  
 CTTTCTCTTC AACAGGTCCA ATCTGACACG GCCCAAACT ATACCATATA CTATTCCATA 720  
 AGAGGTCTCG GAGTTGACCA AGAACCTCGG AATTATTTT ATGTGGAGAG AGACACTGGA 780  
 AACTTGTATT GTACTCGTCC TGTAGATGCT GAGCAGTATG AATCTTTTGA GATAATTGCC 840  
 TTGCAACAA CTCCAGATGG GTATCTCCA GAACCTCCAT TGCCCTTAAT AATCAAAATA 900  
 GAGGATGAAA ATGATAACTA CCAATTTTT ACAGAGAAA CTTATACTTT TACAATTTTT 960  
 GAAATATGCA GAGTGGGCGC TACTGTGGGA CAAGTGTGTG CTACTGACAA AGATGAGCCT 1020  
 GACACGATGC ACACACGCGT GAAGTACTCC ATCATGGGCG AGGTGCCACC ATCACCCACC 1080  
 CTATTTCTTA TGCATCCAAC TACAGGCGTG ATCACCACAA CATCATCTCA GCTAGACAGA 1140  
 GAGTTAATTG ACAAGTACCA GTTGAAAATA AAGTACAAG ACATGGATGG TCAATTTTTT 1200  
 GGTCTACAGA CAACCTCAAC TTGTATCATT AACATTGATG ATGTAATGA CCACTTGCCA 1260  
 ACATTACTCT GTACTCTCTA TGTGACATCA GTGGAAGAAA ATACAGTTGA TGTGGAATC 1320  
 TTACGAGTTA CTGTTGAGGA TAAGGACTTA GTGAATACTG CTAACCTGGAG AGCTAATTAT 1380  
 ACCATTTTAA AGGGCAATGA AATGGCAAT TTTAAATTTG TAACAGATGC CAAAACCAAT 1440  
 GAAGGAGTTC TTGTGTAGT TAAGCCTTTG AATTATGAAG AAAAGCAACA GATGATCTTG 1500  
 CAAATTTGGT TAGTTAATGA AGCTCCATT TCCAGAGAGG CTAGTCCAAG ATCAGCCATG 1560  
 AGCAGACGAA CAGTTACTGT TAATGTAGAA GATCAGATG AGGGCCCTGA GTGTAACCT 1620  
 CCAATACAGA CTGTTCCGAT GAAAGAAAAT AGTGGCATAA GGTATAAGAA ATTAACGTAT 1680  
 AAAGCATATG ACCAGAAAC AAGAAGTAG AGTGGCATAA GGTATAAGAA ATTAACGTAT 1740  
 CCAACAGGCT GGGTCAACAT TGATGAAAAT ACAGGATCAA TCAAGTTTT CAGAAGCTG 1800  
 GATAGAGAGG CAGAGACCAT CAAAATGGC ATATATAATA TTACAGTCTC TGCACTAGAC 1860  
 CAGGAGGAGA GAACATGTAC GGGGACACTG GGCATTATAC TTCAAGACGT GAATGATAAC 1920  
 AGCCCATTTA TACCTAAAAA GACAGTGATC ATCTGCAAA CCAACCATGTC ATCTGCGGAG 1980  
 ATTGTTGCGG TTGATCTCTG TGAGCCTATC CATGGCCAC CTTTGTGACT TAGTCTGGAG 2040  
 AGTTCTACTT CAGAAGTACA GAGAAATGTG AGACTGAAAG CAATTAATGA TACAGCAGCA 2100  
 GGTCTTCTCT ATCAGAAATG TCCTCCATT GGTCTATATG TAGTACCTAT AACAGTGAGA 2160  
 GATAGACTTG GCATGTCTAG TGTCACTTCA TTGGATGTTA CACTGTGTGA CTGCAATACC 2220  
 GAAAATGACT GCACATCTG TGTAGATCCA AGGATTGGCG GTGGAGGAGT ACAACTTGA 2280  
 AAGTGGGCGA TCCTTGCAAT ATTGTTGGCG ATAGCAATTG TCTTTTGCAT CCTGTTTACG 2340  
 CTGGTCTGTG GGGCTTCTGG GACGCTTAAA CAACCAAAAG TAATTTCTGA TGATTTAGCC 2400  
 CAGCAGAAC TAATTGTATC AAACACAGAA GCTCCTGGAG ATGACAAAGT GTATTCTGCG 2460  
 AATGGCTTCA CAACCAAAAC TGTGGGCGCT TCTGCTCAGG GAGTTTGTGG CACCGTGGGA 2520  
 TCAGGAATCA AAAACGAGG TCAGGAGACC ATCGAAATGG TGAAGAGGAG ACACAGAGCC 2580  
 TCGGAATCCT GCGGGGGGGG TGGCCACCAT CACACCTGG ACTCCTGCAG GGGAGGACAC 2640  
 ACGGAGGTGG ACAACTGCGAG ATACACTTAC TCGGAGTGGC ACAGTTTAC TCAGCCCGGT 2700  
 CTTGGTGAAG AATCCATTAG AGGACACACT CTGATTAAAA ATTAACAAT GAAAGAAAGT 2760  
 GTATCTGTGT AATCAAGATG AAAATCACA GCATGCCCAA GACTATGTCC TGACATATAA 2820  
 CTATGAAGGA AGAGGATCGG TGGCTGGGTC TGTAGGTTGT TGCACTGAAC GACAAGAGA 2880  
 AGATGGGCTT GAATTTTGG ATAATTGGA GCCCAATTT AGGACACTAG CAGAAGCATG 2940  
 CATGAAGAGA TGAGTGTGTT CTAATAAGTC TCTGAAAGCC AGTGGCTTTA TGACTTTTAA 3000  
 AAAAATTAC AAACCAAGAA TTTTAAAG CAGAAGATGC TATTGTGGG GGTTTTCTC 3060  
 TCATTATTG GATGGAATCT CTTTGTCAA ATGCACATT ACAGAGAGAC ACTATAACA 3120  
 AGTACAGAAA TTTTCAATT TTTACATT TTTAAATTAC TTATCTTCTA TCCAAGGAGG 3180  
 TCTACAGAGA AATTAAAGTC TGCTTATTT GTTACATTG GGTATAATGA CAACAGCCAA 3240  
 TTTATAGTGC AATAAAATGT AATTAAATCA AGTCTTATT ATAGACTATT TGAAGCACA 3300  
 CCTAATGGAA AATTGTAGAG ACCTTGCTTT AACATTATCT CCAATTAAAT AAGTGTCT 3360  
 GTGGTGTCTG GAAACTGTTG TTTTCTGAA CATCTAAAGT GTGTAGACTG CATTCTGTCT 3420  
 ATTATTTTAT TCTGTAAATG TGACCTTTT ACTGTGCAAA GGGAGATTTC TAGCCAGGCA 3480  
 TTGACTATTA CAATTCATT

Seq ID NO: 619 Protein sequence  
 Protein Accession #: NP\_004940.1

1 11 21 31 41 51  
 MEARPSGSW NGALCRLLLL TLAILIFASD ACKNVTLHVP SKLDAEKLVG RVNLRKCFEA 60



ANLIHSSDPD FOILEDGVSVY TTNILLSSSE KRSPTILLSN TENQEKKKIF VFLEHQTKVL 120  
 KKRHTKRIVL RRAKRRWAPI PCSMLENSLG PFPLELQOVQ SDTAQNTIY YSIRGPGVDQ 180  
 EPRNLFFYVER DTGNLYCTRP VDREQYESFE IIAFATTPDG YTPELPLPLI IKIEDENDNY 240  
 PIPTEETYP TIFENCVRGT TVGQVCATDK DEPDTHMTRL KYSIIGQVPP SPTLFSMHPT 300  
 TGVITTTSSQ LDRELIDKYQ LKIKVQDMG QYFGLQTTST CIINIDVDND HLPFTFRTSY 360  
 VTSVEENTVD VBILRVTVED KDLVNTANWR ANYTILKGNE NGNFKIVTDA KTNBGLVCV 420  
 KPLNVEERQQ MILQIGVVNE APFSREASPR SAMSTATVTV NVEDQDEGPE CNPPIQTVRM 480  
 KENAEVGTTS NGYKAYDPET RSSSGIRYKK LTDPTOWVTI DENTGSIKVF RSLDREAETI 540  
 KNGIYNITVL ASDQGRRTCT GTLGIIILQDV NDNSPFIPKK TVIICKPTMS SAEIVAVDPD 600  
 EPIHGPFPDF SLESSTSEVQ RMWRKAIND TAARLSYQND PPFGSYVVP I TVRDLGMSS 660  
 VTSLDVTLCD CITENDCTHR VDPRIIGGGV QLGKWAIIAI LLGIALLFCI LFTLVCGASG 720  
 TSKQPKVIPD DLAQQLIVS NTEAPGDDKV YSANGFTTQT VGASAQGVCG TVGSGIKNGG 780  
 QETIEMVKGQ HQTSSECRGA GHHTFLDSCR GGHTEVDNCR YTYSEWHSPT QPRLGEESIR 840  
 GHTLIKN

Seq ID NO: 620 DNA sequence  
 Nucleic Acid Accession #: NM\_032545.1  
 Coding sequence: 46..718

1 11 21 31 41 51  
 AAAGTATCT TCAATGCACT AAGAGAAGGA GACTCTCAA CCAAAATGA CCTGGAGGCA 60  
 CCATGTGAGG CTCTCTGTTA CGGTCACTT GGCATTACAG ATCATCAATT TGGGAAACAG 120  
 CTATCAAAGA GAGAAACATA ACGGCGGTAG AGAGGAAGTC ACCAAGGTTG CCACTCAGAA 180  
 GCACCGACAG TCACCGCTCA ACTGGACCTC CAGTCATTTC GGAGAGGTGA CTGGGAGCGC 240  
 CGAGGGCTGG GGGCGGAGG AGCGCTCCC CTACTCCGG GCTTTCGGAG AGGGTGGCTC 300  
 CGCGCGCGCG CTCTGCTGCA GGAAAGCGCG TACCTGCGTG CTGGGAGCT TCTGCGTGTG 360  
 CCGCGCCACG TCCACCGGCG GCTACTGCGA GCATGACCA AGGCGCAGTG AATGCGGCGC 420  
 CCTGGAGCAC GGAGCCTGGA CCTCCGCGC CTGCCACCTC TGCAGGTGCA TCTTCGGGGC 480  
 CTGCACTGCG CTCCCTCTCC AGACGCTGTA CGCTGTGAC CGAAAGACT TCCTGGGCTC 540  
 CCAAGCTCAC GGGCGAGCG CGGGGGGCGC GCCAGCGCTG CTACTCTGC TGCCCTGCGC 600  
 ACTCTGTCAC CGCTCTCTGC GCCCGGATGC GCCCGGCGC CCTCGTCCC TGCTCCCTTC 660  
 CGTCTCCAG CGGAGCGGCG GCCCTGCGG AAGGCGGGA CTGGGCGAT GCCTTTAATT 720  
 TTCTATGTTG TAAATAATAG ATGTGTTTAG TTTACCGTAA GCTGAAGCAC TGGGTGAATA 780  
 TTTTATTGGA GTAATAATA TTTTCATGAA AGCGCCAAAA AAAAAAAAAA AAAAAAAAAA 840  
 AAAAAA

Seq ID NO: 621 Protein sequence  
 Protein Accession #: NP\_115934.1

1 11 21 31 41 51  
 MTWRHVRLL FTVSLALQII NLGNSYQREK HNGGREEVTK VATQKRRQSP LNWTSSEHFE 60  
 VTGSAEGWGP EEPLPYSRAP GEGASARPRC CRNGGTCVLG SPCVCPAHFT GRYCEHDQRR 120  
 SECCALEHGA WTLRACHLCR CIFGALHCLP LQTPDRCDPK DPLASHAGP SAGGAPSLLL 180  
 LLPCALLERL LRPDAPAPRP SLVPSVLQRE RRPCGRPLGL HRL

Seq ID NO: 622 DNA sequence  
 Nucleic Acid Accession #: FGENESH predicted  
 Coding sequence: 1..390

1 11 21 31 41 51  
 ATGAGGTTCA GTGTCTCAGG CATGAGGACC GACTACCCCA GGAGTGTGCT GGCTCCTGCT 60  
 TATGTGTGAG TCTGTCTPCT CCTCTTGTG CCAAGGGAAG TCATCGCTCC CGCTGGCTCA 120  
 GAACCATGGC TGTGCCAGCC GGCACCCAGG TGTGGAGACA AGATCTACAA CCCCTTGGAG 180  
 CAGTGTGTTT ACAAATGACG CATCGTGTCC CTGAGCGAGA CCGGCCAATG TGCTCCCCC 240  
 TGCACTTCTT GGCCTCTGCT TGAGCTCTGC TGTCTTGATT CCTTTGGCCT CACAAACGAT 300  
 TTTGTTGTGA AGCTGAAGGT TCAGGCTGTG AATTCCAGT GCCACTCATC TCCCATCTCC 360  
 AGTAAATGTG AAGAGGCGG GATATGTTAG

Seq ID NO: 623 Protein sequence  
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51  
 MRFSVSGMRT DYPRSVLAPA YVSVCLLLLC PREVIAPAGS EPWLQCPAPR CGDKIYNPLE 60  
 QCCYNDAIVS LSETRQCGPP CTFWPCFELC CLDSFGLTND FVVKLVQGV NSQCHSSPIS 120  
 SKCERGRIC

Seq ID NO: 624 DNA sequence  
 Nucleic Acid Accession #: M18728.1  
 Coding sequence: 51..1085

1 11 21 31 41 51  
 GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60  
 CCTCAGCCCC TCCCTGCAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCCTCAC 120  
 TTCTAACTT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180  
 ATGTGCGAGA GGGGAAGGAG GTTCTTCTAC TCGCCACCAA CTGCCCCAG AATCGTATTG 240  
 GTTACAGCTG GTACAAAGCG GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300  
 TAGGAACCTA ACAAGCTACC CCAGGCGCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360  
 ATGCATCCCT GCTGATCCAG AACGTCACCC AGAATGACAC AGSATTCTAT ACCCTACAAG 420  
 TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGAGC 480  
 TGCCCAAGCC CTCCTCTCTC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540  
 CCTTCACTCG TGAACCTGAG GTTCAGAAAC CAACCTACCT GTGGTGGGTA AATGGTCAGA 600  
 GCCTCCCGGT CAGTCCACAG CTGCAGCTGT CCAATGGCAA CATGACCCCT ACTCTACTCA 660

GCGTCAAAAG GAAAGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720  
 ACCGCAGTGA CCCAGTCACC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCC 780  
 CCTCAAGGCG CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA 840  
 5 ACCCACCTGC ACAGTACTCT TGGTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900  
 TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960  
 CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020  
 TCCTCTCAGC TGTGGCCACC GTCCGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080  
 TATAGCAGCC CTGGTGTATT TTGGATATTT CAGGAAGACT GGCAGATTGG ACCGAGCCCT 1140  
 10 GAATTCCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGCTCTG 1200  
 CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAACTC AATGAAAAAT TAAAGGGAAA 1260  
 ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACTAGAGA CAGTCAAACT 1320  
 GCAAAACCATG GTGAGAAATT GACGACTTCA CACTATGAC AGCTTTTCCC AAGATGTCAA 1380  
 AACAGAGCTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCCT TGCTTATGCC 1440  
 15 TGCCTCTTTC AGCTTGGCAG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500  
 GGGTAACCTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTTAC ATAAAAAAG 1560  
 AGATCCCTTA GTGACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTC 1620  
 AAATGTACAG TGGTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTCTT CACTCCCTGT 1680  
 TTTAATTCAA CCCAGCCATG CAATGCCAAA TAATAGAAAT GCTCCCTACC AGCTGAACAG 1740  
 20 GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800  
 GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860  
 CTGACTCATT CTTTATCTTA TTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920  
 CTCTTGGTAT TACCTCCTCA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980  
 25 CTCTAAAAGC TTTAAATGTC TGCAATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040  
 GGCTGGAATT ERCAAACTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100  
 ATAAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGTTACAC 2160  
 TCTCACTTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220  
 GTTAAGGAAG AGAGTAGATC CAATTAATAA AAATTAATAA CAATTAATAA AAAAAAAGA 2280  
 ACACAGGAGA TTCCAGCTCA CTTGAGTTAG CATAATACAG AAGTCCCTCT TACTTTAAT 2340  
 30 TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCAA TGTATTTATT TCTGTGGTTC 2400  
 TGTTCCTTGG TTCCAAATTG ACAAAACCCA CTGTCTCTGT ATTGTATTGC CCAGGGGGAG 2460  
 CTATCACTGT ACTGTAGAG TGGTGTCTGT TTAATTCATA AATCACAAAT AAAAGCCAAT 2520  
 TAGCTCTATA ACT

Seq ID NO: 625 Protein sequence  
 Protein Accession #: AAA59907.1

1 11 21 31 41 51  
 MGPPSAPPCR LHPVWKEVLL TASLLTFWNP PTTAKLTIES TPFNVABEKE VILLAHNLPO 60  
 40 NRIGYSWYKG ERVDGNSLIV GYVIGTQAT PGPAYSGRET IYPNASLLIQ NVQTQNDTGFY 120  
 TLQVTKSDLV NEBAATGQPHV YPELKPSPIS SNNSNPFVEDK DAVAPTCEPE VQNTTYLWVV 180  
 NQGSFLVSPR LQLSNGNMTL TLLSVKRND A GSYECBIQNP ASANRSDEVT LNVLYGPDVP 240  
 TISPSKANYR PGENLNLSCQ AASNPPAQYS WFINGTFQOS TQELFIPNIT VNNSGSYMCQ 300  
 45 AHSNATGLNR TTVTMITVSG SAPVLSAVAT VGITIGVLAR VALI

Seq ID NO: 626 DNA sequence  
 Nucleic Acid Accession #: M18728.1  
 Coding sequence: 1355..1657

1 11 21 31 41 51  
 GGAGCTCAAG CTCTCTTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60  
 CCTCAGCCCC TCCTCTCAGA TTGCATGTCC OCTGGAAGGA GGTCTGTCTC ACAGCCTCAC 120  
 55 TTCTAACTCT CTGGAACCCA CCCACCACTG CCAAGCTCAC TATTGAATCC ACGCCATTCA 180  
 ATGTCCGAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA OCTGCCCCAG AATCGTATTG 240  
 GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300  
 TAGGAATCA ACAAGCTACC CCAGGCCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360  
 ATGCTATCCT GCTGATCCAG AACGTCACCC AGAATGACAC AGGATTCTAT ACCCTACAAG 420  
 60 TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGGAGC 480  
 TGCCCAAGCC CTCCATCTCC AGCAACAACT CCAACCCCGT GGAGGACAAG GATGCTGTGG 540  
 CTTTCACTGT TGAACCTGAG GTTCAGAACA CAACCTACCT GTGGTGGGTA AATGTCAGTA 600  
 GCCTCCCGGT CAGTCCCAGG GTGCAGTGT CCAATGGCAA CATGACCCTC ACTCTACTCA 660  
 65 GCGTCAAAAG GAAAGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720  
 ACCGCAGTGA CCCAGTCACC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCC 780  
 CCTCAAGGCG CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA 840  
 ACCCACCTGC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900  
 TCTTTATCCC CAACATCACT GTGAATAATA GCGGATCCTA TATGTGCCAA GCCCATAACT 960  
 CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGCTCCTG 1020  
 70 TCCTCTCAGC TGTGGCCACC GTCCGCATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080  
 TATAGCAGCC CTGGTGTATT TTGGATATTT CAGGAAGACT GGCAGATTGG ACCGAGCCCT 1140  
 GAATTCCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGCTCTG 1200  
 CTCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAACTC AATGAAAAAT TAAAGGGAAA 1260  
 75 ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACTAGAGA CAGTCAAACT 1320  
 GCAAAACCATG GTGAGAAATT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380  
 AACAGAGCTC CTCATCATGA TAAGGCTCTT ACCCCCTTTT AATTTGTCCT TGCTTATGCC 1440  
 TGCCTCTTTC GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAAGAAG TAGCTTCAGA 1500  
 GGGTAACCTA ACAGAGTGTC AGATCTATCT TGTCAATCCC AACGTTTTAC ATAAAAAAG 1560  
 80 AGATCCCTTA GTGACCCAG TGACTGACAT TAGCAGCATC TTTAACACAG CCGTGTGTTC 1620  
 AAATGTACAG TGGTCTTTT CAGAGTTGGA CTCTAGACT CACCTGTCTT CACTCCCTGT 1680  
 TTTAATTCAA CCCAGCCATG CAATGCCAAA TAATAGAAAT GCTCCCTACC AGCTGAACAG 1740  
 GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800  
 SCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860  
 85 CTGACTCATT CTTTATCTTA TTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920  
 CTCTTGGTAT TACCTCCTCA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980  
 TCTTAAAAGC TTTAAATGTC TGCAATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040  
 GGCTGGAATT ACAAAACCTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100  
 ATAAAAGCCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTGGTCACAC 2160

TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220  
 GTTAAGGAAG AGATAGATGC CAATTAAAAA AAATTAAAC CAATTAAAAA AAAAAAAGA 2280  
 ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTC TACTTTAACT 2340  
 TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCA TGTATTATT TCTGTGGTTC 2400  
 TGTTCCTTG TTCCAATTG ACAAACCCA CTGTTCTGT ATTGTATTGC CCAGGGGGAG 2460  
 CTATCACTGT ACTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAT AAAAGCCAAT 2520  
 TAGCTCTATA ACT

Seq ID NO: 627 Protein sequence  
 Protein Accession #: AAA59908.1

1 11 21 31 41 51  
 MDSFSQDVKT RLLIMIRLLP PPNLSLLMPA SPANQDDAVI SISQEVASEG NLTECQIYLV 60  
 NPNVLHKIRD PLVHPVTDIS SIFNTAVCSN VQMSFSELDP

Seq ID NO: 628 DNA sequence  
 Nucleic Acid Accession #: M18728.1  
 Coding sequence: 2370..2501

1 11 21 31 41 51  
 GGAGCTCAAG CTCCTCTACA AAGAGGTGGA CAGAGAAGAC AGCAGAGACC ATGGGACCCC 60  
 CCTCAGCCCC TCCCTGCAGA TTGCATGTCC CCTGGAAGGA GGTCTGCTC ACAGCTCAC 120  
 TTCTAACCTT CTGGAACCCA CCCACCCTG CCAAGCTCAC TATTGAATCC AGGCCATTCA 180  
 ATGTGCGAGA GGGGAAGGAG GTTCTTCTAC TCGCCACAA CCTGCCCCAG AATCGTATTG 240  
 GTTACAGCTG GTACAAAGGC GAAAGAGTGG ATGGCAACAG TCTAATTGTA GGATATGTAA 300  
 TAGGAATCTA ACAGACTACC CCAGGGCCCG CATACAGTGG TCGAGAGACA ATATACCCCA 360  
 ATGCATCCCT GCTGATCCAG AACGTACACC AGAATGACAC AGGATTCTAT ACCCTACAAG 420  
 TCATAAAGTC AGATCTGTG AATGAAGAAG CAACCGGACA GTTCCATGTA TACCCGAGGC 480  
 TGCCCAAGCC TCCCATCTCC AGCAACAAC CCAACCCCGT GGAGGACAAG GATGCTGTGG 540  
 CCTTCACTG TGAACTCTAG GTTCAGAA CAACCTACCT GTGGTGGGTA AATGGTCAGA 600  
 GCCTCCCGGT CAGTCCAGG CTGCAGCTGT CCAATGGCAA CATGACCCCT ACTCTACTCA 660  
 GCGTCAAAAG GAAAGATGCA GGATCCTATG AATGTGAAAT ACAGAACCCA GCGAGTGCCA 720  
 ACCGCACTGA CCCAGTCACC CTGAATGTCC TCTATGGCCC AGATGTCCCC ACCATTTCCT 780  
 CCTCAAAGGC CAATTACCGT CCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA 840  
 ACCCACTGAC ACAGTACTCT TGGTTTATCA ATGGGACGTT CCAGCAATCC ACACAAGAGC 900  
 TCTTTATCCC CACAGTACTC GTGAATAATA GCGGATCCTA TATGTGCCAA GCCATAACT 960  
 CAGCCACTGG CCTCAATAGG ACCACAGTCA CGATGATCAC AGTCTCTGGA AGTGTCTCTG 1020  
 TCCTCTCAGC TGTGGCCACC GTCCGATCA CGATTGGAGT GCTGGCCAGG GTGGCTCTGA 1080  
 TATAGCAGCC CTGGTGATTT TCGATATTT CAGGAAGACT GGCAGATTGG ACCAGACCTT 1140  
 GAATCTCTCT AGCTCCTCCA ATCCCATTTT ATCCCATGGA ACCACTAAAA ACAAGGTCTG 1200  
 TCTGCTCCT GAAGCCCTAT ATGCTGGAGA TGGACAATC AATGAAAAAT TAAAGGGAAA 1260  
 ACCCTCAGGC CTGAGGTGTG TGCCACTCAG AGACTTCACC TAACTAGAGA CAGTCAAACT 1320  
 GCAAACCATG GTGAGAAAT GACGACTTCA CACTATGGAC AGCTTTTCCC AAGATGTCAA 1380  
 AACAAAGCTC TCATCATGA TAAGGCTCTT ACCCCCTTTT AATTGTCTCT TGCTTATGCC 1440  
 TGCTCTTCT GCTTGGCAGG ATGATGCTGT CATTAGTATT TCACAGAAG TAGCTTCAGA 1500  
 GGGTAACCTA ACAGAGTGTG AGATCTATCT TGTCAATCCC AACGTTTAC ATAAAAAAG 1560  
 AGATCCTTTA GTGCACCCAG TGACTGACAT TAGCAGCATC TTAAACACAG CCGTGTGTTT 1620  
 AATGTACAG TGGTCTTTT CAGAGTTTGA CTCTAGACT CACCTGTTCT CACTCCCTGT 1680  
 TTTAATTCAA CCCAGCCATG CAATGCCAAA TAATAGAATT GCTCCCTACC AGCTGAACAG 1740  
 GGAGGAGTCT GTGCAGTTTC TGACACTTGT TGTGAACAT GGCTAAATAC AATGGGTATC 1800  
 GCTGAGACTA AGTTGTAGAA ATTAACAAAT GTGCTGCTTG GTTAAATGG CTACACTCAT 1860  
 CTGACTCATT CTTTATCTTA TTTTAGTTGG TTTGTATCTT GCCTAAGGTG CGTAGTCCAA 1920  
 CTCTTGTGAT TACCTCTCTA ATAGTCATAC TAGTAGTCAT ACTCCCTGGT GTAGTGTATT 1980  
 CTCTAAAAGC TTTAAATGTC TGCATGCAGC CAGCCATCAA ATAGTGAATG GTCTCTCTTT 2040  
 GGCTGGAATT ACAAATCTCA GAGAAATGTG TCATCAGGAG AACATCATAA CCCATGAAGG 2100  
 ATAAAGAGCC CAAATGGTGG TAACTGATAA TAGCACTAAT GCTTTAAGAT TTTGCTACAC 2160  
 TCTCACCTAG GTGAGCGCAT TGAGCCAGTG GTGCTAAATG CTACATACTC CAACTGAAAT 2220  
 GTTAAGGAAG AGATAGATGC CAATTAAAAA AAATTAAAC CAATTAAAAA AAAAAAAGA 2280  
 ACACAGGAGA TTCCAGTCTA CTGAGTTAG CATAATACAG AAGTCCCTC TACTTTAACT 2340  
 TTTACAAAAA AGTAACCTGA ACTAATCTGA TGTAAACCA TGTATTATT TCTGTGGTTC 2400  
 TGTTCCTTG TTCCAATTG ACAAACCCA CTGTTCTGT ATTGTATTGC CCAGGGGGAG 2460  
 CTATCACTGT ACTGTAGAG TGGTGCTGCT TTAATTCATA AATCACAAT AAAAGCCAAT 2520  
 TAGCTCTATA ACT

Seq ID NO: 629 Protein sequence  
 Protein Accession #: AAA59909.1

1 11 21 31 41 51  
 MLTNVFIQV LFPCSNLTKP TVLVLYCPGG AITVLVENCC FNS

Seq ID NO: 630 DNA sequence  
 Nucleic Acid Accession #: NM\_016639.1  
 Coding sequence: 40..429

1 11 21 31 41 51  
 GCGGCGGGCG CAGACAGCGG CCGGCGCAGG ACCTGCACTA TGGCTCGGGG CTCGCTGCGC 60  
 CGGTTGCTGC GGCTCCTCGT GCTGGGGCTC TGGCTGGCGT TGCTGCGCTC CGTGGCCGGG 120  
 GAGCAAGGCG CAGGCACCGC CCCCTGCTCC CGCGCGAGCT CTTGAGCGSC GGACCTGGAC 180  
 AAGTGATGAG ACTGCGCGTC TTGCGGGGCG CGACCGCACA GCGACTTCTG CTTGGGCTGC 240  
 GCTGCAGCAC CTCTGCGCCC CTTCGCGCTG CTTTGGGCCA TCTTGGGGGG CGCTCTGAGC 300  
 CTGACCTTGG TGCTGGGGCT GCTTCTGCGC TTTTGGTCTT GGAGACGATG CCGCAGGAGA 360  
 GAGAAGTTCA CCACCCCATC AGAGGAGACC GCGGAGAGG GCTGCCAGC TGTGGCGCTG 420

ATCCAGTGC AATGTGCCCC CTGCCAGCCG GGGCTCGCCC ACTCATCATT CATTCTATCCA 480  
 TTCTAGAGCC AGTCTCTGCC TCCCAGACGC GCGCGGAGCC AAGCTCCTCC AACCAACAGG 540  
 GGGGTGGGGG GGGGTGAATC ACCTCTGAGG CCTGGGCCCA GGGTTCAGGG GAACCTTCCA 600  
 AGGTGTCTGG TTGCCCTGCC TCTGGCTCCA GAACAGAAAG GGAGCCTCAC GCTGGCTCAC 660  
 ACAAACACGC TGACACTGAC TAAGGAACCT CAGCATTTGC ACAGGGGAGG GGGGTGCCCT 720  
 CCTTCTTAG GACCTGGGGG CCAGGCTGAC TTGGGGGGCA GACTTGACAC TAGGCCCCAC 780  
 TCACTCAGAT GTCCTGAAAT TCCACCACGG GGGTCAACCT GGGGGGTAG GGACCTATTT 840  
 TTAACACTAG GGGCTGGCCC ACTAGGAGGG CTGGCCCTAA GATACAGACC CCCCACACTC 900  
 CCCAAGCGG GGAGGAGATA TTTATTTGG GGAGAGTTG GAGGGGAGG AGATTATTAT 960  
 AATAAAGAA TCTTTAACTT TAAAAAATA AAAAAA

Seq ID NO: 631 Protein sequence  
 Protein Accession #: NP\_057723.1

1 11 21 31 41 51  
 MARGSLRRLL RLLVLGLWLA LLRSVAGEQA PGTAPCSRGS SWSADLDKCM DCASCRARPH 60  
 SDFCLGCAAA PPAPRLLWP ILGGLSLTF VLGLLSGLFV WRRCRREKF TPIETGTGE 120  
 GCPAVALIQ

Seq ID NO: 632 DNA sequence  
 Nucleic Acid Accession #: NM\_003816.1  
 Coding sequence: 79..2538

25 1 11 21 31 41 51  
 CGGCAGGGTT GGAATATGAT GGAAGAGGCG GAGGTGGAGG GAGCCGAGTG CTGAGAGGAA 60  
 CCTCGGGAAT CGGCCGAGAT GGGGTCTGGC GCGCGCTTTC CCTCGGGGAC CCTTCGTGTC 120  
 CGGTGGTTGC TGTTCCTTGG CCTGGTGGGC CCAGTCTCTG GTGCGGCGCG GCCAGGCTTT 180  
 CAACAGACCT CACATCTTTC TTCTTATGAA ATTATAACTC CTGTGAGATT AACTAGAGAA 240  
 AGAAGAGAA GGCCTAGGCC CTATTCAAAA CAAGTATCTT ATGTTATTCA GGCTGAAGGA 300  
 AAAGAGCATA TTATTCACCT GGAAGGAAC AAAGACCTTT TGCTGAAGA TTTTGGTGT 360  
 TATACTTACA ACAAGGAAGG GACTTTAATC ACTGACCATC CCAATATACA GAATCATGT 420  
 CATTATCGGG GCTATGTGGA GGGAGTTTCA AATTATCCA TTGCTCTTAG CGACTGTTTT 480  
 GGACTCAGAG GATTGTGCA TTTAGAGAA GCGAGTTATG GGATTGAACC CCTGCAGAAC 540  
 AGCTCTCATT TTGAGCACAT CATTATCGA ATGGATGATG TCTACAAAGA GCCTCTGAAA 600  
 TGTGGAGTTT CCAACAAGGA TATAGAGAAA GAACTGCAA AGGATGAAGA GGAAGAGCCT 660  
 CCCAGCATGA CTCAGTACT TCGAAGAAGA AGAGCTGTCT TGCCACAGAC CCGTATGTG 720  
 GAGCTGTTCA TTGTCTGAGA CAAGGAAAGG TATGACATGA TGGGAAGAAA TCAGACTGCT 780  
 GTGAGAGAAG AGATGATTCT CCTGGCAAC TACTTGGATA GTATGTATAT TATGTTAAAT 840  
 ATTGGAATTG TGTAGTTGG ACTGGAGATT TGGACCAATG GAAACCTGAT CAACATAGTT 900  
 GGGGTGCTGT GTGATGTGCT GGGGAACCTC GTGCAGTGGC GGGAAAGATT TCTTATCACA 960  
 CGTCGGAGAC ATGACAGTGC ACAGCTAGTT CTAAAGAAAG GTTTTGGTGG AACTGCAGGA 1020  
 ATGGCATTTC TGGGAACAGT GTGTTCAAGG AGCCACGAG GCGGGATTAA TGTGTTTGA 1080  
 CAAATCACTG TGGAGACATT TGCTTCCATT GTTGCTCATG AATTGGGTCA TAATCTTGA 1140  
 ATGAATCAGC ATGATGGGAG AGATTGTTC TGTGGAGCAA AGAGCTGCAT CATGAATTCA 1200  
 GGAGCATCGG GTTCCAGAAA CTTTAGCAGT TGCAGTGCAG AGGACTTTGA GAAGTTAACT 1260  
 TTAAATAAAG GAGGAAAGT CCTTCTTAAT ATTCCAAAGC CTGATGAAGC CTATAGTGCT 1320  
 CCTCTCTGTG GTAATAAGTT GGTGGACCGT GGGGAAGAGT GTGACTGTGG TACTCCAAAG 1380  
 GAATGTGAAT TGGACCCCTG CTGCGAAGGA AGTACCTGTA AGCTTAAATC ATTTGCTGAG 1440  
 TGTGCATATG GTGACTGTTG TAAAGACTGT CGGTTCTTTC CAGGAGGTAC TTTATGCCGA 1500  
 GGAAAAACCA GTTCTGATGA TGTTCAGAG TACTGCAATG GTTCTTCTCA GTTCTGTGAG 1560  
 CCAGATGTTT TTATTCAGAA TGGATATCCT TGCCAGAATA ACAAAGCCTA TTGCTACAA 1620  
 GGCATGTGCC AGTATTATGA TGCTCAATGT CAAGTCATCT TTGGCTCAA AGCCAGGCT 1680  
 GCCCCCAAGG ATGTTTTCAT TGAAGTGAAT TCTAAAGGTG ACAGATTGG CAATTGTGGT 1740  
 TTCTCTGGCA ATGAATACAA GAAGTGTGCC ACTGGGAATG CTTTGTGTGG AAAGCTTCAG 1800  
 TGTGAGATAG TACAGAGAT ACCTGTATTT GGAATTGTGC CTGCTATTAT TCAACCGCT 1860  
 AGTCGAGGCA CCAATGTGTG GGGTGTGGAT TTCCAGCTAG GATCAGATGT TCCAGATCCT 1920  
 GGGATGTTA ACAGAGGCAC AAAATGTGGT GCTGGAAAGA TCTGTAGAAA CTTCAGTGT 1980  
 GTAGATGCTT CTGTTCTGAA TTATGACTGT GATGTTTACA AAAAGTGTCA TGGACATGGG 2040  
 GTATGTAATA GCAATAGAAA TTGTCACGTG GAAAAAGGCT GGGCTCCCCC AAATTGTGAG 2100  
 ACTAAAGGAT ACAGGAGGAG TGTGGACAGT GGAACCTACAT AATATGAAT GAATCTGCA 2160  
 TTGAGGGACG GACTTCTGGT CTCTTCTTC CTAATTGTTT CCCTTATTGT CTGTGCTATT 2220  
 TTTATCTTCA TCAAGAGGGA TCAACTGTGG AGAAGCTACT TCAGAAAGAA GAGATCACAA 2280  
 ACATATGAGT CAGATGGCAA AAATCAAGCA AACCTTCTA GACAGCGGG GAGTGTTCCT 2340  
 CGACATGTTT CTCCAGTGAC ACCCTCCAGA GAAGTTCTTA TATATGAAA CAGATTGCA 2400  
 GTACCAACCT ATGCAGCCAA GCAACCTCAG CAGTTCCTAT CAAGGCCACC TCCACCACAA 2460  
 CGAAAGATAT CATCTCAGGG AAACCTAATT CCTGCCGCTC CTGCTCTGAC ACCTCCTTTA 2520  
 TATAGTTCCC TCACCTGATT TTTTAACTT TCTTTTTCGA AATGCTCTCA GGGAACTGAG 2580  
 CTAATACCTT TTTTCTTCT TGATGTTTTC TTGAAAGGCC TTTCTGTGTC AACTATGAAT 2640  
 GAAAAACAAA CACCACAAA CAGACTTAC TAACACAGAA AAACAGAAAC TGAGTGTGAG 2700  
 AGTTGTGAAA TACAAGGAAA TGCAGTAAAG CCAGGGAATT TACATAACA TTTCCGTTTC 2760  
 CATCATTGAA TAAGTCTTAT TCAGTATCG GTGAGGTAA TGCATAATC ATGGATTTTT 2820  
 TGAACATGTT ATTGCAGTGA TTCTCAAATT AACTGTATTG GTGTAAGATT TTTGTCAAT 2880  
 AGTGTTTAAG TGTATTCTG AATTTCTAC CTTAGTTATC ATTAATGTAG TTCTCATTT 2940  
 AACATGTGAT AATCTAATAC CTGTGAAAC TGACTAATCA GCTGCCAATA ATATCTAATA 3000  
 TTTTTCATCA TGCAGGAATT AATAATCATC ATACTTAGA ATCTGTCTG TCACTCACTA 3060  
 CATGAATAAG CAAATATTGT CTTCAAAAGA ATGCACAAGA ACCACAATA AGATGTCTA 3120  
 TTATTTGAAA AGTACAAAAT ATACTAAAAG AGTGTGTGTG TATTCACGCA GTTACTCGCT 3180  
 TCCATTTTCA TCACTTTTCA ACTATAGGTA ATAACCTTAA GAGAAATTAA TTTAATATTA 3240  
 GAATTTCTAT TATGAATCAT GTGAAGCAT GACATTCGTT CACAATAGCA CTATTTTAAA 3300  
 TAAATTATAA GCTTTAAGGT ACAGAGTATT TAATAGATCT AATCAATAT GTTGATTCAT 3360  
 GGCTATAATA AAGCAGGAGC AATTATAAAA TCTTCAATCA ATTGAACCTT TACAAAACCA 3420  
 CTTGAGATTA TCAATGAGCAC TTTAAATCT GAACCTTCAA AGCTTGCTAT TAAATCATTT 3480  
 AGAATGTTTA CATTACTTAA GGTGTGCTGG GTCATGTAAA ATATTAGACA CTAATATTTT 3540  
 CATAGAAATT AGGCTGGAGA AAGAAGGAG AAATGGTTTT CTTAAATACC TACAAAAG 3600  
 TTACTGTGTT ATCTATGAGT TATCATCTTA GCTGTGTTAA AAATGAATTT TTACTATGGC 3660

AGATATGGTA TGGATCGTAA AATTTTAAGC ACTAAAAATT TTTTCATAAC CTTTCATAAT 3720  
 AAAGTTTAAT AATAGGTATA TTAACCTGAAT TTCATTAGTT TTTTAAAGT GTTTTGGTT 3780  
 TGTGTATATA TACATATACA AATACAACAT TTACAATAAA TAAATACTT GAAATTCTCA 3840  
 AAAAAAAAAA AAAAAAAAAA AAAAA

Seq ID NO: 633 Protein sequence  
 Protein Accession #: NP\_003807.1

1 11 21 31 41 51  
 | | | | |  
 MGSARFPSSG TLRVRWLLL GLVGPVLGAA RPPGQOTSHL SSYEIITPWR LTRERREAPR 60  
 PYSKQSVYVI QAEGKEHIIH LERNKDLLPE DFVVYTYNKE GTLITDHPNI QNHCHYRGYV 120  
 EGVHNSIAL SDGFLGRLL HLENASYGIE PLQNSHFEH IYRMDVYK EPLKCGVSNK 180  
 DIEKETAKDE ESEPPSMTQL LRRRAVLPO TRYVELFIV DKERYDMGR NQTAVREEMI 240  
 LLANVLDSEY IMLNTRIVLV GLEIWTNGNL INIVGGAGDV LGNFVQWREK FLITRRRHDS 300  
 AQLVLKRGFG GTAGMAFVGT VCSRSAGGI NVFQGITVET FASIVAHLEL HNLGMNDDG 360  
 RDCSCQAKSC IMNSGASGR NFSSCSAEDF EKLTLNKGGL CLNIPKPEE AYSAPSCKNK 420  
 LVDAGEECDC GTPKCELDLP CCEGSTCKLK SPAECAYGDC CKDCRFLPGG TLCRGKTSK 480  
 DVPEYCNSSS QFCQPDVFIQ NGYPCQNNKA YCYNGMCQYY DAQCQVIFGS KAKAAPKDCF 540  
 IEVNSKGRDF GNCQPSGNEY KKCATGNALC GKLQCEVQBE IPVFGIVPAI IQTPSRGTC 600  
 WGVDPQLGSD VPDPMVNEG TKCGAGKICR NFQCVDAVL NYDCDVQKKC HGHGVCNSNK 660  
 NCHCENGWAP FNCETKYGG SVDSGPTYNE MNTALRDGLL VFFFLIVPLI VCAIFIFIKR 720  
 DQLWRSYFRK KRSQTYESDG KNQANPSRQP GSVPRHVSFV TPREVPIYA NRPVAPTYAA 780  
 KQPQFPSPRP PPPQPKVSSQ GNLIAPAP APPLYSSLT

Seq ID NO: 634 DNA sequence  
 Nucleic Acid Accession #: NM\_002091.1  
 Coding sequence: 56..503

1 11 21 31 41 51  
 | | | | |  
 AGTCTCTGCT CTTCACAGCC TCTCCGGGCG GCTCCAAAGG CTTCCTGTCG GGACCATGCG 60  
 CGGAGTGAAG CTCCCGCTGG TCTGCTGGC GCTGGTCTCT TGCTTAGCGC CCGGGGGGCG 120  
 AGCGGTCCCG CTGCGTGGCG GCGGAGGGAC CGTGTCTGAC AAGATGTACC CGCGCGGCAA 180  
 CCACTGGGCG GTGGGGCACT TAATGGGGAA AAAGAGCACA GGGGAGTCTT CTTCTGTTTC 240  
 TGAGAGAGGG AGCCTGAAGC AGCAGCTGAG AGAGTACATC AGGTGGGAAG AAGCTGCAAG 300  
 GAATTTGCTG GTCTCTATAG AAGCAAGGGA GAACAGAAAC CACCAAGCCAC CTCAACCCAA 360  
 GGCTTGGGCG AATCAGCAGC CTTCGTGGGA TTCAGAGGAT AGCAGCAACT TCAAGATGT 420  
 AGGTTCAAAA GGCAAGTGT GTAGACTCTC TGCTCCAGGT TCTCAACGTG AAGGAAGGAA 480  
 CCCCAGCTG ACCAGCAAT GATAATGATG GCCTCTCTCA AAAGAGAAAA ACAAACCCC 540  
 TAAGAGACTG AGTTCTGCAA GCATCAGTTC TACGGATCAT CAACAAGATT TCCTTGTGCA 600  
 AATATTGTA CTATTCTGTA TCTTTCATCC TTGACTAAAT TGTGATTTT CAAGCAGCAT 660  
 CTTCTGGTTT AAATCTGTTT GCTGTGAACA ATTGTGAAA AGAGTCTTCC AATTAATGCT 720  
 TTTTATATC TAGGCTACCT GTTGTTAGA TTCAAGGCCC CGAGCTGTTA CCATTCACAA 780  
 TAAAGCTTA AACACAT

Seq ID NO: 635 Protein sequence  
 Protein Accession #: NP\_002082.1

1 11 21 31 41 51  
 | | | | |  
 MRGSELPLVL LALVLCAPR GRAVPLPAG GTVLTQMYPR GNHNAVGHLM GHKSTGESS 60  
 VSEKSLKQQL LREYRWEEA ARNLLGLIEA KENRNHQPPQ PKALGNQPS WSESDSSNPK 120  
 DVGSKGVGR LSAPGSQREG RNPQLNQ

Seq ID NO: 636 DNA sequence  
 Nucleic Acid Accession #: NM\_016522.1  
 Coding sequence: 265..1299

1 11 21 31 41 51  
 | | | | |  
 GCGGAAGCAG CGAGGAGGGA GCCCCTTTG GCGGTCTTCC GTGGAACCGG TTTTCCGAGG 60  
 CTGGCAAAAG CCGAGGCTGG ATTTGGGGA GGAATATTAG ACTCGGAGGA GTCTGCGCGC 120  
 TTTTCTCTC CCGCGCCTC CCGGTGCGCG GCGGTTCACC GCTCAGTCCC CGCGCTCGCT 180  
 CCGCACCCCA CCCACTTCTT GTGCTGCGCC GGGGGGCGTG TGCGGTGCGG CTGCGGAGT 240  
 TCGGGGAAGT TGTGGCTGTC GAGAAATGGG GTCTGTGGGT ACCTGTTCCT GCCCTGGAAG 300  
 TGCTCTGTTG TGTGTCTCT CAGGCTGCTG TTCTTGTAC CCACAGGAGT GCCCGTGGCG 360  
 AGCGGAGATG CCACCTTCCC CAAAGCTATG GACAACTGA CCGTCCGGCA GGGGAGAGC 420  
 GCCACCTCA GTTGCACTAT TGACAACCGG GTCAACCGGG TGGCTGGCT AAACCGCAGC 480  
 ACCATCTCT ATGCTGGGAA TGACAAGTGG TGCTGGATC CTGCGTGGT CTTCTGAGC 540  
 AACACCCAAA CGCAGTACAG CATCGAGATC CAGAACGTGG ATGTGTATGA CGAGGGCCCT 600  
 TACACCTGCT CGGTGCAGAG AGACAACAC CCAAGACCT CTAGGGTCCA CCTCATTTGT 660  
 CAAGTATCTC CCAAAATTGT AGAGATTCT TCAGATATCT CCATTAATGA AGGGAACAAT 720  
 ATTAGCCTCA CTGTCATAGC AACTGGTAGA CCAGAGCCTA CGTTACTTGT GAGACATC 780  
 TCTCCCAAAG CGTTGGCTT TGTAGTGAA GACGAATACT TGGAAATTCA GGGCATCACC 840  
 CCGGAACAGT CAGGGGACTA CGAGTGCAGT GCCTCCAATG ACGTGGCGCG GCCCGTGGTA 900  
 CGGAGAGTAA AGGTCACGT GAACATATCA CCATACATTT CAGAAGCCAA GGGTACAGGT 960  
 GTCCCGGTGG GACAAAAGGG GACACTGCAG TGTGAAGCCT CAGCAGTCCC CTCAGCAGAA 1020  
 TTCCAGTGGT ACAAAGATGA CAAAAGACTG ATTGAAGGAA AGAAGGGGT GAAAGTGAA 1080  
 AACAGACCTT TCTCTCAA ACTCATCTT TCAATGTCT CTGAACATGA CTATGGGAAC 1140  
 TACACTTGGG TGGCTCCAA CAAGCTGCGC CACACCAATG CCAGCATCAT GCTATTGGT 1200  
 CCAGGCGCGG TCAGCGAGGT GAGCAACGCG ACGTCAAGGA GGGCAGGCTG CTTCTGCTG 1260  
 CTGCTCTTTC TGGTCTTGCA CCGTCTTCT AAATTTTGT GTGAGTGCCA CTTCCACCAC 1320  
 CGGGAAGGCG TGCCGCCACC ACCACCACA ACACAACAGC AATGGCAACA CCGACAGCAA 1380  
 CCAATCAGAT ATATCAAAT GAAATTAGAA GAAACACAGC CTCATGGGAC AGAAATTGTA 1440  
 GGGAGGGGAA CAAAGAAATC TTTGGGGGGA AAAGAGTTT AAAAAAGAAA TTGAAATG 1500  
 CCTTCAGAT ATTTAGGTAC AATGGAGTTT TCTTTTCCCA AACGGGAAGA ACACAGCACA 1560

WO 02/086443

CCCCGCTTGG ACCCACTGCA AGCTGCATCG TGCAACCTCT TTGGTGCCAG TGTGGGCAAG 1620  
 GGCTCAGCCT CTCTGCCAC AGACTGCCCC CAGCTGGAAC ATTCTGGAGC TGGCCATCCC 1680  
 AAATTCAATC AGTCCATAGA GACGAACAGA ATGAGACCTT CCGGCCCAAG CGTGGCGCTT 1740  
 CCGGCCCAAG CGTGGCGCTG CCGGCACCTT GGTGACTGT GCCACCAAG CGTGTGTGT 1800  
 5 GAAACGTGAA ATAAAGAG CAAAAA AAAAAA

Seq ID NO: 637 Protein sequence  
 Protein Accession #: NP\_057606.1

10 1 11 21 31 41 51  
 MGVCGYLFLP WKCLVVVSLR LLPLVPTGVP VRSGDATFPK AMDNVTVRQG ESATLRCTID 60  
 NRVTRVAVLN RSTILYAGND KNCIDPRVVL LSNITQYYSI EIQNVDVYDE GPYTCVQTD 120  
 15 NHPKTSRVHL IVQVSPKIVE ISSDISINEG NNISLTICAT GRPEPTVWTR HISPKAUGPV 180  
 SEDEYLEIQG ITRRQSGDYE CSASNDVAAP VVRRVKVTYN YPPYISEAKG TGVPVQKGT 240  
 LQCEASAVPS AEFQMYKDDK RLIEGKGVK VENRPFLSKL IFFNVSEHDY GNYTCVASNK 300  
 LGHTNASIML PGPGAVSEVS NGTSRRAGCV WLLPLLVHL LKLF

Seq ID NO: 638 DNA sequence  
 Nucleic Acid Accession #: NM\_012261.1  
 Coding sequence: 203..1045

25 1 11 21 31 41 51  
 GATTTCCTCT GCCAGCAGCT GTCGGTGCGG CGCTCGACAC CGAGTCCTAG CTAGGCGCTC 60  
 ACAGAAATAG CGCTCCCTCC CTCCCTCTTC TCTGTCCGCC GCCTCTCGCT CACCCCGGCC 120  
 CACTCCAGCG GCGACTTTGA GGGATTCCCT CTCTGCGCGC CTCTGCAGCA GCACAGCCGG 180  
 CCTCATTCGG GGCACCTGGA GTATGATCT CCAAGGAAGA GGGTCCCCA GCATCGACAG 240  
 30 ACTTCAGATT CTCTGTATGT CCACTAACC TGAAAAAGAT ATATTGTGG TGGGGAAGA 300  
 GGAATAATCT TCAGGCTTT CAGAGTTTG AGCCAAATTT ATTGTACCTT ATGATGTGTG 420  
 TGGGAGCAGC TGCTCATGG CAGAGTTTG AGCCAAATTT ATTGTACCTT ATGATGTGTG 480  
 TGAGGTGAAG GGCCTGTGT GCCACAGCCA GTCCGAGCTG CAAGTGTCTT GGGTGGATCG 540  
 35 GGCAGCAAC TAGTAGATC TGATCACAGA ACAGCCGAT ATCGCATTA CCGCGGAGC 600  
 TGAGGTGAAG GGCCTGTGT GCCACAGCCA GTCCGAGCTG CAAGTGTCTT GGGTGGATCG 660  
 CGCATATGCA CTCAAATGC TCTTTGTAAA GGAAGCCAC AACATGTCCA AGGACCTGA 720  
 GGCAGCTGG AGGCTGAGCA AAGTGCAGT TGCTACGAC TCCTCGGAGA AAACCCACTT 780  
 CAAAGACGCA GTCAAGCTG GGAAGCACAC AGCCAACTCG CACCACTCT CTGCTTGGT 840  
 CACCCCGCT GGGAACTCT ATGAGTGTCA AGCTCAACAA ACCATTTCAC TGGCCTCTAG 900  
 40 TGATCCGAG AAGACGCTCA CCATGATCT GTCTGCGCTC CACATCCAAC CTTTGTGACT 960  
 TATCTCAGAT TTTGTCTTCA GTGAAGAGCA TAAATGCCCA GTGGATGAGC GGGAGCAACT 1020  
 GGAAGAAACC TTGCCCCGTA TTTTGGGGCT CATCTTGGGC CTGCTCATCA TGGTAACACT 1080  
 CGCGATTAC CAGCTCCACC ACAAATGAC GTAGAGGCC GTTAGGCAG CACCCCTAT TCCTGTCTCC 1140  
 ATCCAGTAT AAGCAGATG GCTAGAGGCC GTTAGGCAG CACCCCTAT TCCTGTCTCC 1200  
 CCAACTGGAT CAGGTAGAAC AACAAAGCA CTTTCCATC TTGTACAGA GATACACAA 1260  
 45 CATAGCTACA ATCAACAGG CTTGGGTATC TGAGGCTTGC TTGGCTTGTG TCCATGCTTA 1320  
 AACCACGSA AGGGGAGAC TCTTTCGGAT TTGTAGGCTG AAATGGCAAT TATTCTCTCC 1380  
 ATGCTGGGGA GAGGGGAGG AGGGCTCAG ACAGCTTTCG TGCTCATGGT GCGCTTGGCT 1440  
 TGACTCTCCA AAGAGCAATA AATGCCACTT GAGCTGTAT CTGGCCCAA AGTTTAGGA 1500  
 TTGAAACAT GCTTCTTGA GAGGAAACC CCTTTAGGTT CAGAAGAATA TGGGTGCTT 1560  
 50 TGCTCCCTTG GACACAGCTG GCTTATCTTA TACAGTGTG AATGCACACA GAATACAA 1620  
 TCATGCTCCC TGCAGCAAGA CCCCAGAAAG TGATTCATGC TTCTGGCTGG CATTCTGCAT 1680  
 GTTTAGTATG TGCTTGGGA ATGTTTCACT GCTACCCGCA TCCAGCGACT CGACACCAAG 1740  
 AAAACGACTA ATGTAATAT CGAGAGTGT TTGGACTTCT TCCTGTGCCA GGTCCAAGT  
 GGGGACCTG AAGATCAAT CTGTGTGAGT CTGTTTTC AATGAAATA AACACACTA  
 TTCTCTGGC

Seq ID NO: 639 Protein sequence  
 Protein Accession #: NP\_036393.1

60 1 11 21 31 41 51  
 MDLQGRGVPS IDRLRVLLML FHMAQIMAB QEVNLSGLS TNPEKDIFV RENGTTCLMA 60  
 EFPAKFIIVP DWASNYVDL ITEQADIALT RGAEVKRCG HSQSEIQVFW VDRAYALML 120  
 FVKESHNMRS GPEATWRLSK VQFVDSSEK THFKDAVSAG KHTANSHLS ALVTPAGKSY 180  
 65 ECQAQQTISL ASSDPQKTVT MILSAVHIQ FDIISDFVFS EERKCPVDER EQLEETLPLI 240  
 LGLILGLVIM VTLAIYVHR KMTANQVQIP RDRSQYKMG

Seq ID NO: 640 DNA sequence  
 Nucleic Acid Accession #: NM\_002993.1  
 Coding sequence: 64..408

70 1 11 21 31 41 51  
 GGCACAGGCC AGTCTCCGGC CCTCCACCCA GCTCAGGAAC CCGGGAACCC TCTCTTGACC 60  
 ACTATAGGCC TCCCTGCCAG CCGCGCGGCC CGTGTCCCGG GTCCCTCGGG CTCCTTGTGC 120  
 75 GCGCTGCTCG CGCTGCTGCT CCGCTGACG CCGCGCGGCC CCCTCGCCAG CGCTGGTCTC 180  
 GTCTCTGCTG TGCTGACAGA GCTGCGTTGC ACTGTCTTAC GCGTTACGCT GAGAGTAAAC 240  
 CCCAAACAGA TTGTTAACT CGAGGTGTC CCCGAGGCC CGCAGTGCTC CAAGGTGGAA 300  
 GTGGTAGCCT CCGTGAAGAA CGGGAAGCAA GTTTGTCTGG ACCCGGAAGC CCCTTTCTA 360  
 80 AAGAAAGTCA TCCAGAAAT TTTGACAGT GGAACACAGA AAACTGAGT AACAAAAAG 420  
 ACCATGCATC ATAAATATGC CCACTCTTCA CGGAGCAGT TTTCTGAGA TCCCTGACC 480  
 CAGTAAGAAT AAGAGGAAG GGTGGTTTT TTTCCATTT CTACATGAT TCCCTACTTT 540  
 GAAGAGTGTG GGGGAAGCC TACGCTTCTC CCGTGAAGTT ACAGCTCAG TAATGAAGTA 600  
 CTAATATAGT ATTTCCACTA TTACTGTTA TTTTACCTGA TAAGTTATTG AACCTTTGG 660  
 85 CAATTGACCA TATTGTGAGC AAGAAATCAC TGGTTATTAG TCTTCAATG AATATTGAAT 720  
 TGAAGATAAC TATTGTATT CTATCATACA TCTCTTAAAG TCTTACCGAA AAGGCTGTGG 780  
 ATTTCTGATG GAAATAATGT TTTATTAGT TGCTGTGAG GGAGGTATCC TGTGTCTCT 840  
 ACTCACTCTT CTCATAAAT AGGAAATATT TTAGTCTCTG TTTCTTGGG AATATGTTAC 900

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85

Seq ID NO: 641 Protein sequence  
Protein Accession #: NP\_002984.1

1 11 21 31 41 51  
| | | | | |  
MSLPSSRAAR VPGPSGSLCA LLALLLLLT PGLASAGPV SAVLTELRCT CLRVTLRVNP 60  
KTIGKLQVFP AGPQCCKVEV VASLRNGKQV CLDPEAPFLK KVIQKILDSG NKKN

Seq ID NO: 642 DNA sequence  
Nucleic Acid Accession #: NM\_013271.1  
Coding sequence: 27..809

1 11 21 31 41 51  
| | | | | |  
TCCGGAGCCA GGCTGCTGG GGCAGCATGG CGGGGTGCGC GCTGCTCTGG GGGCGCGGG 60  
CCGGGGGGGT CGGCCCTTTT GTGCTGCTGC TGCTGCGCCT GTTTCGCGCG CCCCCGCGC 120  
CTGCGCGCG GCGGTAAAG GAACCCCGCG GCCTAAGCGC AGCGTCTCGC CCCTTGCGTG 180  
AGACTGGCGC TCCTCGCGCG TTCGCGCGGT CAGTGCCCGC AGGTGAGGCG GCGGGGGCGG 240  
TGCAGGAGCT GCGCGCGCGC CTGGCGCATC TGCTGGAGGC CGAACGTCAG GAGCGGGCGC 300  
GGCGCGAGGC GCAGGAGGCT GAGGATCAGC AGGCGCGCGT CTGGCGCAG CTGCTGCGCG 360  
TCTGGGGCGC CCCCCGCAAC TCTGATCGCG CTCTGGGCTT GACGACGAC CCGGACGCGC 420  
CTGCAGCGCA GCTGCTGCG GCTCTGCTGC GCGCGCGCCT TGACCCGCGC GCCCTAGCAG 480  
CCGAGCTTGT CCGCGCGCGC GTCCCGCGCG CGGCGCTCGC ACCCGGCGCC CGGCTCTAGC 540  
ACGAGCGGCG CCGCGCGCGC GATGCTGAGG AGGCAGGCGA CGAGACACCC GACGTGGACC 600  
CCGAGCTGTT GAGGTACTTG CTGGGACGGA TTCTTCGCGG AAGCGCGGAC TCGAGGGGG 660  
TGGCAGCCCC CGCGCGCCTC CGCGGTGCGC CGACCCAGCA TGTGGGCTCT GAGCTGCCCC 720  
CTGAGGGCGT GCTGGGGCGC CTGCTGCGTG TGAAACGCGT AGAGACCCCG GCGCCCCAGG 780  
TGCTTGACCG CGGCTCTTG CCACCGTGAG CACTGCGCGG ATCCGCTGCA CCCTGGGACC 840  
CAGAAGTGCT CCGCGCATTC CGCCACGAG ACTTCTCCCC GCCAGCACGT CCAGAGCAAC 900  
TTACCCGCGC CAGCCAGCCC TCTCACCAGA GGATCCCTAC CCCCTGGCCC ACAATAACAT 960  
GATCTGAGC

Seq ID NO: 643 Protein sequence  
Protein Accession #: NP\_037403.1

1 11 21 31 41 51  
| | | | | |  
MAGSPLLWGP RAGGVGLLVL LLLGLFRPPP ALCARPVKEP RGLSAASPPL AETGAPRRFR 60  
RSVPGEAAG AVQELARALA HLLERQER ARARAQERAD QQARVLAQLL RVWGAPRNSD 120  
PALGLDDDPD AFPAQLARAL LRARLDPAAL AAQLVPAFVP AAALRFRPPV YDDGPAGPDA 180  
EAGADETFDV DPELLRYLLG RILAGSADSE GVAAPRRLRR AADHDVGSSEL PPEGVLGALL 240  
RVKRLTEPAP QVPARRLLPP

Seq ID NO: 644 DNA sequence  
Nucleic Acid Accession #: NM\_002214  
Coding sequence: 681..2990

60  
65  
70  
75  
80  
85

1 11 21 31 41 51  
| | | | | |  
CCAGAGCCG CCTCCCCCTG TTGCTGGCAT CCGAGCTTC CTCCCTTGCC AGCCAGGAGC 60  
CTGCCGACTT GTCTTTGCCG GCTGCTCGCG AGACGGGGCT GCAAGCTGCG AACTAATGGT 120  
GTTCGCTTCC CTGCCACCTT GTGGAAGCAA CTGCGCTGAT TGATGCGCCA CAGACTTTT 180  
TCCCTCGAG CTGCGCGCGG TACCTCCCA CAGATCCAGC ATCACCAGT GAATGTACAT 240  
TAGGGTGGTT TCCCGCCAG CTTGCGGCTT TGTTTGGGTT TGATTGTGTT TGGCTCTTCG 300  
CTAAGCTGAT TTATGCAGCA GAAGCCCGAC CGGCTGAGGA GAAACAAAG CTCTTTTCTT 360  
TGTCCCGGAG CAGGCTGCGG AGCCCTTGCA GAGCCCTCTC TCCAGTCCGC GCCGGGCGCT 420  
TGGCGGTGGA AGGAGGTGCT TCTCGCGGAG ACCCGGGGAC CGCGCGTGCC GAGCGGGGAG 480  
GGCGGTAGGG GCCCTGAGAT GCGGAGCGGT GCCCGGGGCC GCTTACCTGC ACCGCTTGCT 540  
CCGAGCCGCG GGGTCCGCTT GCTAGGCTTG CCGAAACGCT CTTAGCGACA CTGCGCCGCG 600  
GGCCCGGAGG TGCCCGGGA GCGGAGCGCC GCGTCCGGA GGCAGCCAGG CGCGGGGCGC 660  
GGGGCGGGCT GTTTTGCAAT ATGTGCGGCT CGGCCCTGGC TTTTITTAAC GCTGCATTG 720  
TCTGCTGCA AAACGACCGG CGAGGTCCCG CCTGTTCTCT CTGGGCGAGC TGGGTGTTT 780  
CACTTGTTCT TGGAGTGGGC CAAGGTGAAG ACAATAGATG TGCACTTTCA AATGCAGCAT 840  
CCTGTGCCAG GTGCTTGCG CTGGGTCCAG AATGTGATG GTGTGTTCAA GAGGATTTC 900  
TTTCAAGTGG ATCAGAAGT GAACGTTGTG ATATTGTTTC CAATTAATA AGCAAAGGCT 960  
GCTCAGTTGA TTCAATAGAA TACCATCTG TGCACTGTTT AATACCCACT GAAATGAAA 1020  
TTAATACCCA GGTGACACCA GGAGAAGTGT CTATCCAGCT GCGTCCAGGA GCCGAGGCTA 1080  
ATTTTATGCT GAAAGTTTCT CCTCTGAAGA AATATCTGTT GGATCTTTAT TATCTGTTG 1140  
ATGCTCTGAG ATCAATGAC AATAATATAG AAAAATTAAA TTCGTGTTGA AACGATTAT 1200  
CTAGAAAAAT GGCATTTTTC TCCCTGACT TTGCTCTGG ATTTGCTCA TACGTTGATA 1260  
AAACAGTTTC ACCATACATT AGCATCCACC CCGAAAGGAT TCATAATCAA TGCAGTGACT 1320  
ACAAATTAGA CTGCATGCTT CCCCATTGAT ACATCCATGT GCTGTCTTTG ACAGAGAACA 1380  
TCACGTAGTT TGAGAACGCA GTTCATAGAC AGAAGATCTC TGGAAACATA GATACACCA 1440  
AAGGAGGTTT TGAGCCCATG CTTAGGCGAG CTGCTGTGTA AAGTCATATC GATGGGCGAA 1500  
AAGAGGCTAA AAGATTGCTG CTGGTGATGA CAGATCAGAC GTCTCATCTC GCTCTTGATA 1560

GCAAATTGGC AGGCATAGTG GTGCCCAATG ACGGAAACTG TCATCTGAAA AACAACTGCT 1620  
ACGTCAATATC GACACCATG GAACACCCCT CACTAGGCCA ACTTTCAGAG AAATTAATAG 1680  
ACAACAACAT TAATGTGATC TTGCGATTC AAGGAAACAA ATTTTATGCG TATAAGGATC 1740  
TTCTACCCCT CTGCGCAAGC ACCATTGCTG GTGAAATAGA ATCAAAGGCT GCAAACTTCA 1800  
ATAATTTGGT AGTGGAAAGC TATCAGAAAGC TCATTTCAGA AGTGAAAGIT CAGGTGGAAA 1860  
ACCAGGTACA AGGCATCTAT TTAAACATTA CCGCATCTG TCCAGATGGG TCCAGAAAGC 1920  
CAGGCATGGA AGGATGCGAG AACGTGACGA GCAATGATGA AGTTCITTTT AATGTAACAG 1980  
TTACAAATGAA AAAATGTGAT GTACAGGAG GAAAAAATA TGCAATAATC AAACCTATTG 2040  
GTTTTAATGA AACCGCTAAA ATTCATATAC ACAGAAACTG CAGCTGTCAG TGTGAGGACA 2100  
ACAGAGGACC TAAAGGAAAG TGTGTAGATG AAACCTTTCT AGATTCCAAG TGTTCAGT 2160  
GTGATGAGAA TAAATGTGAT TTGATGAGG ATCAGTTTTC TTCTGAGAGT TGCAAGTCA 2220  
ACAAGGATCA GCCTGTTTGC AGTGGTCGAG GAGTTTGTGT TTGTGGGAAA TGTTCATGTC 2280  
ACAAATTAAG GCTTGGAAAA GTGTATGGA AATACTGTGA AAAGGATGAC TTTTCTTGT 2340  
CATATACCA TGAATATCTG TGTGCTGGGC ATGGAGAGTG TGAAGCAGGC AGATGCCAAT 2400  
GCTTCAGTGG CTGGGAAGGT GATCGATGCC AGTGCCCTTC AGCAGCAGCC CAGCACTGTG 2460  
TCAATTCAA GGGCCAAGTG TGCAGTGGAA GAGGCACTG TGTGTGTGGA AGGTGTGAGT 2520  
GCACCGATCC CAGGAGCATC GCGCGCTTCT GTGAACACTG CCCCACCTGT TATACAGCT 2580  
GCAAGGAAAA CTGGAATTTG ATGCAATGCC TTCACCTCA CAATTGTCT CAGGCTATAC 2640  
TTGATCAGTG CAAACCTCA TGTGCTCTCA TGGAAACAA GCATTATGTC GACCAACTT 2700  
CAGAATGTTT CTCCAGCCCA AGCTACTTGA GAATATTTT CATCATTTT ATAGTTACAT 2760  
TCTTGTATGG GTTGCTTAAA GTCTGATCA TTAGACAGGT GATACTACAA TGGAAATAGT 2820  
ATAAAATTA GTCTCTTCA GATTACAGAG TGTGAGCTC AAAAAAGGAT AAGTTGATTC 2880  
TGCAAGTGT TTGCAAGA GCACTCAGT ACCGACGTGA GAAAGCTGAA GAAATAAAA 2940  
TGGATATCAG CAAATTAAT GCTCATGAAA CTTTCAGTGT CAACTTCTAA AAAAAAGATT 3000  
TTAAACACTT AATGGGAAAC TGGAAATTTT AATAATGCT CCTAAGATT AATAATTTAA 3060  
AAGTCAAGG AGGAGACAAA TTGCTCAGG TCATGCCAGT TGCTGGTGT ACACCTCGAAC 3120  
GAAGACTGAC AAGTATCTTC ATCATGATGT GACTCACATA GCTGCTGACT TTTTCAGAGA 3180  
AAAATGTGTC TTAATCTAGT TTGAGACTAG TGTGTTGTA GCACCTTACT GTAATATATA 3240  
ACTTATTTAG ATCAGCATAG AATGTAGATC CTCTGAAGAG CACTGATTAC ACTTTACAGG 3300  
TACCTGTTAT CCCTAGCTT CCCAGAGAGA ACAATGCTGT GAGAGAGTT AGCATTTGT 3360  
CACTACAAG GTACAGTAAT CCCTGCACTG GACATGTGAG GAAAAAATA ATCTGGCAAG 3420  
TATATTCTAA GGTGTCACAA CACTTCAACA GTTGGTGGTT GAATAGACAA GAACAGCTAG 3480  
ATGAATTAAT GATTCTGTT TCACTCTTC AAGAGGTGAA CAGATACAAC CTTAATCTTA 3540  
AAAGATTAT GCTTTTAAA GTGTGTAGT TTATGCTAT GTGTTTATG TTTGCTTAT 3600  
TTGCAAGAT GGATCAAT TCCAGCATTC TCTCTCTTT GCCTTTATGT TTTGTTTCT 3660  
TTTTACAGG ATAAGTTAT GTATGTCACA GATGACTGGA TTAATTAAGT GCTAAGTTAC 3720  
TACTGCCATA AAAAACTAAT AATACAATGT CACTTATCA GAATACTAGT TTTAAAGCT 3780  
GAATGTTAA

Seq ID NO: 645 Protein sequence  
Protein Accession #: NP\_002205

1 11 21 31 41 51  
MGSSALAFPT AAFVCLQNDR RGPASFLWAA WVFSVLVLGL QGEDNRCASS NAASCARCLA 60  
LGPEQHCWCV EDPISSGSSRS ERCDIVSNLI SKGCSVDSIE YPSVHVIIPT ENHINTQVTP 120  
GEVSIQLRPG AEANFMKLVH PLKKYPVDLY YLVDVSASMH NNIEKLNSVG NDLSRKMFAFF 180  
SRDFRLGFGS YVDKTVSPYI SIHPERIHQ CSDYNLDQMP PHGYIEVLST TENITEPEKA 240  
VHRQKISGNI DTPEGGFDAM LQAAVCESHI GWRKEAKRLI LVMTDQTSIL ALDSKLQIV 300  
VPNDGNCHLK NNIVVKSTTM EHPSLGLQSE KLIDNNINVI FAVQGRQPHN YKDLLPLPLG 360  
TIAGIESKA ANLNLNVEEA YOKLISEVKV QVENQVQGIY FNITAICPDG SRKPGMEGCR 420  
NVISNDEVLF NVVTVMKCD VTKGKNYAI KPIGFNETAK IHIHRNCSQ CEDNRGPKGK 480  
CVDETFLDSK CFQCDENKCH FDEQDFSSS CKSHKDPVC SGRGVCCVCG CSCHKIKLKG 540  
VYGKYCEKDD FSCPYHGNL CAGHGECEAG RCQCFSGWEG DRQCPSSAAA OHCNVNSKGV 600  
CSGRGTVCVG RCECTDPRSI GRPCEHCPTC YTACKRNWC MQCLHPHNL OAILDQCKTS 660  
CALMEQQRHV DOTSECFSSP SYLRIFFIIF IVTFILGLLK VLIIRQVILQ WNSNLIKSSS 720  
DYRVSASKKD KLILQSVCTR AVTYRREKPE EIKMDISKLN AHETFRCNF

Seq ID NO: 646 DNA sequence  
Nucleic Acid Accession #: NM\_003318.1  
Coding sequence: 1..2574

1 11 21 31 41 51  
ATGGAATCCG AGGATTAAAG TGGCAGAGAA TTGACAATTG ATTCCATAAT GAACAAAGTG 60  
AGAGACATTA AAAATAAGTT TAAAAATGAA GACCTTACTG ATGAATAAG CTGGAATAAA 120  
ATTTCTGCTG ATACTACAGA TAACTCGGGA ACTGTTAACC AAATTATGAT GATGGCAATC 180  
AACCAGAGG ACTGTTGAG TTTGTTGCTC AAACTAGAGA AAAACAGTGT TCCGCTAAGT 240  
GATGCTCTTT TAAATAAATT GATTGGTCTG TACAGTCAAG CAATTGAAGC GCTTCCOCCA 300  
GATAAATATG GCCAAATGA GAGTTTGTCT AGAATTCAAG TGAGATTGTC TGAATTAAAA 360  
GCTATTCAAG AGCCAGATGA TGCAAGTGAC TACTTTCAAA TGGCCAGAGC AAACCTGCAAG 420  
AAATTTGCTT TTGTTTATAT ATCTTTTGCA CAATTGAAC TGTCAACAG TAATGTCAAA 480  
AAAAGTAAC AACTTCTTCA AAAAGCTGTA GAACCTGGAG CAGTACCCT AGAATGCTG 540  
GAAATTGCCC TGGGGAATTT AAACCTCFAA AAAAGCAGC TGCTTTTCA GAGGAGAAAG 600  
AAGAAATTTA CAGCATCTAC GGTATTAACT GCCAAGAAT CATTTTCCGG TTCACTTGGG 660  
CATTACAGA ATAGGAACAA CAGTTGTGAT TCCAGAGGAC AGACTACTAA AGCCAGGTTT 720  
TTATATGGAG AGAATATGCC ACCACAAGT GCAGAAATAG GTTACCGGAA TTCATTGAGA 780  
CAAACTAACA AACTAAACA GTCATGCCCA TTGGAAGAG TCCAGTTAA CCTTCTAAT 840  
AGCCAGATT GTGATGTGAA GACAGATGAT TCAGTTGTAC CTGTTTAT GAAAGACAA 900  
ACCTCTAGAT CAGAAATGCC AGATTGTTT GTGCTGGAT CTAACCAAG TGGAAATGAT 960  
TCTGTGAAT TAAGAAATTT AAGTCTGTT CAAATAGTC ATTTCAAGGA ACCTCTGTT 1020  
TCAGATGAAA AGAGTTCTGA ACTTATTATT ACTGATTCAA TAACCTGAA GAATAAAAG 1080  
GAATCAAGTC TTCTAGCTAA ATTAGAAGAA ACTAAAGAGT ATCAAGAAC AGAGGTCCA 1140  
GAGAGTAACC AGAAACAGTG GCAATCTAAG AGAAGTCTAG AGTGATTAA CCAGATCTCT 1200  
GCTGCTCTT CAATCACTG GCAGATTCCG GAGTTAGCCC GAAAGTTAA TACAGAGCAG 1260  
AAACATACCA CTTTGTAGCA ACCTGTCTT TCAGTTTCAA AACAGTCAAC ACCAATATCA 1320  
ACATCTAAAT GGTTTGACCC AAAATCTATT TGTAGACAC CAAGCAGCAA TACCTTGGAT 1380



GATTACATGA GCTGTTTGTAG AACTCCAGTT GTAAAGAATG ACTTTCACC TGCTTGTGAG 1440  
 TTGTCAACAC CTTATGGCCA ACCTGCCTGT TTCCAGCAGC AACAGCATCA AATACTTGCC 1500  
 ACTCCACTTC AAAATTTCAC GGTTTTAGCA TCTTCTTCAG CAAATGAATG CATTTCCGTT 1560  
 AAAGGAGAGAA TTTATTCCAT TTTAAGCAG ATAGGAAGTG GAGGTTCAAG CAAGGTATTT 1620  
 CAGGTGTTAA ATGAAAAGAA ACAGATATAT GCTATAAAAT ATGTGAACCT AGAAGAAGCA 1680  
 GATAACCAAA CTCTTGATAG TTACCGGAAC GAAATAGCTT ATTTGAATAA ACTACAACAA 1740  
 CACAGTGATA AGATCATCCG ACTTTATGAT TATGAAATCA CGGACCAGTA CATCTACATG 1800  
 GTAATGGAGT GTGGAATAT TGATCTTAAT AGTTGGCTTA AAAAGAAAAA ATCCATTGAT 1860  
 CCATGGGAAC GCAGAGTTA CTGGAATAAT ATGTTAGAGG CAGTTCAACAC AATCCATCAA 1920  
 CATGGCATTG TTCACAGTGA TCTTAAACCA GCTAACTTTC TGATAGTTGA TGGATGCTA 1980  
 AAGCTAATTG ATTTTGGGAT TGCAACCAAA ATGCAACCAG ATACAACAAG TGTGTGTTAA 2040  
 GATTCTCAGG TTGGCAGCT TAATTATATG CCACCAGAAG CAATCAAAGA TATGCTCTCC 2100  
 TCCAGAGAGA ATGGGAATC TAAGTCAAAG ATAAGCCCCA AAGTGATGT TTGTCCTTA 2160  
 GGATGTATT TGTACTATAT GACTTACGGG AAAACACCAT TTCAGCAGAT AATTAATCAG 2220  
 ATTTCTAAAT TACATGCCAT AATTGATCCT AATCATGAAA TTGAATTTCC CGATATTCCA 2280  
 GAGAAGATC TCAAGATGT GTTAAAGTGT TGTGTTAAAA GGGACCCAAA ACAGAGGATA 2340  
 TCCATTCTCG AGCTCCCTGC TCATCCCTAT GTTCAAATTC AAACCTATCC AGTTAACCAA 2400  
 ATGGCCAAAG GAACCACTGA AGAAATGAAA TATGTTCTGG GCCAATTGT TGGTCTGAAT 2460  
 TCTCTTAATC CCATTTTGAA AGCTGCTAAA ACTTTATATG AACACTATAG TGGTGGTGAA 2520  
 AGTCATAATT CTTCATCCTC CAAGACTTTT GAAAAAATAA GGGGAAAAAA ATGA

Seq ID NO: 647 Protein sequence  
 Protein Accession #: NP\_003309.1

1 11 21 31 41 51  
 MESEDLSGRE LTIDSIMNKV RDIKNKFKNE DLTDELSLNK ISADTTDSNG TVNQIMMMAN 60  
 NPZDWLSLLL KLEKNSVPLS DALLNKLIGR YSQAIKALPP DKYQGNESFA RIQVRFALK 120  
 AIQEPDDARD YQMARANCK KPAFVHISFA QFELSQGNVK KSKQLLQKAV ERGAVPLEML 180  
 EIALRNLMNQ KQQLLSEEEK KNLASSTVLT AQESFSGSLG HLQNRNNSCD SRGQTTKARF 240  
 LYGENMPQPD AETGYRNLKP QTNKTKQSCP FGRVFNLLN SPDCDVKTDD SVVPCFMKRG 300  
 TTSRECRDLV VPGSKPSNGD SCLELNKLSV QNSHFKEPLV SDEKSEELII TDSITLKNKT 360  
 ESSLLAKLEE TKEYQEPEVP ESNQKQWQSK RKSECIQNPN AASSNHWQIP ELARKVNTAQ 420  
 KHTTFBQVPV SVSKQSPPI S TSKWDFPKSI CKTPSSNTLD DYMSCRPTVP VKNDPPACQ 480  
 LSTPYGQAPC FQQQHQHILA TPLQNLQVLA SSSANECISV KGRIYSILKQ IGSGGSSKVF 540  
 QVLNEKKQIY AIKYVNLBEE DNQTLDSYRN EIAVYLNKLQO HSDKIIRLYD YEITDQYIYM 600  
 VMCEGNIDLN SWLKKKSID PWERKSYWKN MLEAVHTIHQ HGIVHSDLPK ANPLIVDGML 660  
 KLIDFGIANQ MQPDTTSVVK DSQVGTNYNM PPEALKDMSS SRENGKSKSK ISPKSDVWSL 720  
 GCILYMYTYG KTFPQQIINQ ISKLHAIIDP NHEIEFPDIP EKDLQDVLKC CLKRDPKQRI 780  
 SIPELLAHFY VQIQTHPVNQ MAKGTTEEMK YVLGQLVGLN SPNSILKAAK TLYEHYSGGE 840  
 SHNSSSSKTF RKKRKGK

Seq ID NO: 648 DNA sequence  
 Nucleic Acid Accession #: NM\_015507  
 Coding sequence: 241..1902

1 11 21 31 41 51  
 CGGCAGAGGA GCCTCGGCCA GGCTAGCCAG GCGCGCCCCA GCCCTCCCCC AGGCCGCGAG 60  
 CGGCCCTGCC GCGGTGCTCG GCCTCCCTTC CCAGACTGCA GGGACAGCAC CCGGTAATG 120  
 CGAGTGGAGC GGAGGACCCG AGCGGCTGAG GAGAGAGGAG GCGGGCGCTT AGCTGCTACG 180  
 GGGTCCGCCC GCGCGCTTCC CGAGGGGGGC TCAGAGAGGAG GAAGGAGGAC CCGTCCGAGA 240  
 ATGCCCTGTC CTCTGAGCTG TCGCTCCCG CTGCTGCTCT CCGGGTGGC AGGTGTTTC 300  
 GGGAAACGGG CCAATGCAAG GCATCAGCGG TTGTAGCAT CCGCACGTCA GCCTGGGTC 360  
 TGTCACTATG GAATAAACT GGCCTGCTGC TACGGCTGGA GAAGAAACAG CAAGGGAGTC 420  
 TGTGAAGCTA CATGCGAACC TGGATGTAAG TTTGGTGAGT GCGTGGGACC AAACAAATGC 480  
 AGATGCTTTC CAGGATACAC CGGGAACACC TGCACTCAAG ATGTGAATGA GTGTGGAATG 540  
 AAACCCCGGC CATGCAACA CAGATGTGTG AATACACAGC GAAGCTACAA GTGCTTTTGC 600  
 CTCAGTGCCC ACATGCTCAT CCGCAGTGCT ACGTGTGTGA ACTCTAGGAC ATGTGCCATG 660  
 ATAAACTGTC AGTACAGCTG TGAAGACACA GAAGAAGGGC CACAGTGCCT GTGTCCATCC 720  
 TCAGGACTCC GCCTGGCCCC AAATGGAAGA GACTGTCTAG ATATTGATGA ATGTGCTCT 780  
 GGTAAAGTCA TCTGTCCCTA CAATCGAAGA TGTGTGAACA CATTGGAAG CTACTACTGC 840  
 AAATGTACCA TTGGTTTCEA ACTGCAATAT ATCAGTGGAC GATATGACTG TATAGATATA 900  
 AATGAATGTA CTATGGATAG CCATACGTGC AGCCACCATG CCAATTGCTT CAATACCCAA 960  
 GGGTCTTCA AGTGTAAATG CAAGCAGGGA TATAAAGSCA ATGGACTTCG GTGTTCTGCT 1020  
 ATCCCTGAAA ATTCTGTGAA GGAAGTCTTC AGAGCACCTG GTACCATCAA AGACAGAATC 1080  
 AAGAAGTTCG TTGCTCAACA AAACAGCATG AAAAAGAGGG CAAAAATTAA AAATGTTACC 1140  
 CCAGAACCCA CCAGGACTCC TACCCCTAAG GTGAACCTTC AGCCCTTCAA CTATGAAGAG 1200  
 ATAGTTTCCA GAGGCGGGAA CTCTCATGGA GGTAAAAAAG GGAATGAAGA GAAATGAAA 1260  
 GAGGGGCTTG AGGATGAGAA AAGAGAAGAG AAAGCCCTGA AGAATGACAT AGAGGAGCGA 1320  
 AGCCTGCCAG GAGATGTGTT TTTCCCTAAG GTGAATGAAG CAGGTGAATT CGGCCTGATT 1380  
 CTGGTCCAAA GGAAGAGCGT AACTTCCAAA CTGGAACATA AAGATTTAAA TATCTCGGTT 1440  
 GACTGCAGCT TCAATCATGG GATCTGTGAC TGGAAACAGG ATAGAGAAGA TGATTGTGAC 1500  
 TGGATCTCTG CTGATCGAGA TAATGCTATT GGCTTCTATA TGGCAGTTCC GGCCTTGGCA 1560  
 GGTCAAGAAG AAGACATTGG CCGATTGAAA CTTCTCTCTA CTGACCTGCA ACCCCAAAGC 1620  
 AACTTCTGTT TGCTCTTTGA TTACCGGCTG GCCGAGACA AAGTCGGGAA ACTTCGAGTG 1680  
 TTTGTGAAA ACAGTAACAA TGCCCTGGCA TGGGAGAAGA CCACGAGTGA GGATGAAAAG 1740  
 TGGAGACAG GGAATAATCA GTTGATACAA GGAACCTGAT CTACCAAAAG CATCATTTT 1800  
 GAAGCAGAAC GTGGCAAGGG CAAAACCGGC GAAATCGCAG TGGATGGCGT CTTGCTTGT 1860  
 TCAGGCTTAT GTCCAGATAG CCTTTATCT GTGGATGACT GAATGTTACT ATCTTTATAT 1920  
 TTGACTTGTG ATGTCAGTTC CTGTTTCTT TTGATATTGC ATCATAGGAC CTCTGGCATT 1980  
 TTAGAATTAC TAGCTGAAA ATTGTAATGT ACCAACAGAA ATATTATTGT AAGATGCCTT 2040  
 TCTGTATATA GATATGCCAA TATTTGCTTT AAATATCATA TCACTGTATC TTCTCAGTCA 2100  
 TTTCTGAATC TTTCCACATT ATATTATAAA ATATGGAAT GTCACTTTAT CTCCCTCTCT 2160  
 CAGTATATCT GATTGTATA AGTAAGTTGA TGAGCTTCTC TCTACAACAT TTCTAGAAAA 2220  
 TAGAAAAAAA AGCAGAGAGA AATGTTTAACT TGTTTGACTC TTATGATACT TCTTGGAAAC 2280  
 TATGACATCA AAGATAGACT TTTGCCTAAG TGGCTTAGCT GGGTCTTTCA TAGCCAAACT 2340

TGTATATTTA AATTCTTTGT AATAATAATA TCCAAATCAT CAAAAAATAA AAAAAAAA

Seq ID NO: 649 Protein sequence  
Protein Accession #: NP\_056322

1	11	21	31	41	51	
MPLPWSLALP	LLLSWVAGGF	GNAASARHHG	LLASARQPGV	CHYGTKLACC	YGWRRNSKGV	60
CEATCEPGCK	FGCEVGPNC	RCFPGYTGT	CSQDVNECGM	KPRPCQHRV	NTHGSYKCF	120
LSGHMLMPDA	TCVNSRTCAM	INQVSCEDT	EEGPQCLCPS	SGLRLAPNGR	DCLDIDBCAS	180
GKVICPNRR	CVNTFGSYIC	KCHIGFELQY	ISGRYDCIDI	NECTWDSHTC	SHHANCPTQ	240
GSFKCKCQGG	YKGNGLRCSA	IPENSVKELV	RAPGTIKDRI	KKLLAHKNSM	KKKAKIKNVT	300
PEPTRTPPK	VNLQPFNYEE	IVSRGNGSHG	GKKGNEEMK	EGLEDEKREE	KALKNDIEER	360
SLRGDVFFPK	VNEAGEFGLI	LVQRKALTSK	LEHKDLNISV	DCSPNHGICD	WKQDREDDFD	420
WNPAARDNAI	GPYMAVPLAI	GHKKDILRLK	LLLPDLQPOS	NFCLLFDPYRL	AGDKVGLKRV	480
FVKNSNNALA	WEKTTSEDEK	WKTGKIQLYQ	GTDATKSIIF	EAERGKGTG	ELAVDGVLLV	540
SGLCPDLSLS	VDD					

Seq ID NO: 650 DNA sequence  
Nucleic Acid Accession #: NM\_003506.1  
Coding sequence: 259..2379

1	11	21	31	41	51	
GCAGCTCCAG	TCCCGGACGC	AACCCCGGAG	CCGTCTCAGG	TCCCTGGGGG	GAACGGTGGG	60
TTAGACGGGG	ACGGGAAGGG	ACAGCGGCTT	TGACCGGCC	CCCGAGTAAT	TGACCCAGGA	120
CTCATTTTCA	GGAAAGCCTG	AAATAGTAA	AAATAGTGAA	ATGAGGAATT	TGAACATTTT	180
ATCTTTGGAT	GGGATCTTTC	TGAGGATGCA	AAGAGTGATT	CATCCAAGCC	ATGTGGTAAA	240
ATCAGGAATT	TGAAGAAAA	GGAGATGTTT	ACATTTTGTG	TGACGTGTAT	TTTCTACCC	300
CTCCTAAGAG	GGCAGAGTCT	CTTCACTGTG	GAACCAATTA	CTGTTCAGG	ATGTATGAAA	360
ATGGCTTACA	ACATGACGTT	TTTCCCTAAT	CTGATGGGTC	ATTATGACCA	GAGTATTGCC	420
GCGGTGGAAA	TGGAGCATTT	TCTTCTCTC	GCAAATCTGG	AATGTTTACC	AAACATTGAA	480
ACTTTCTCT	GCAAAGCATT	TGTACCAACC	TGCATAGAAC	AAATTCATGT	GGTTCCACCT	540
TGTCTGTAAC	TTTGTGAGAA	AGTATATTCT	GATTGCAAAA	AATTAATTGA	CACITTTGGG	600
ATCCGATGGC	CTGAGGAGCT	TGAATGTGAC	AGATTACAAT	ACTGTGATGA	GACTGTTCTT	660
GTAACCTTTG	ATCCACACAC	AGAATTTCTT	GGTCTTCAGA	AGAAAACAGA	ACAAGTCCAA	720
AGAGACATTG	GATTTTGTGT	TCCAAGGCAT	CTTAAGACTT	CTGGGGGACA	AGGATATAAG	780
TTTCTGGGAA	TGACCAAGTG	TGCGCTTCCA	TGCCCAACAA	TGTATTTTAA	AAGTGTGAG	840
CTAGAGTTTG	CAAAAAGTTT	TATTGGAACA	GTTTCAATAT	TTTGTCTTTG	TGCAACTCTG	900
TTCAACTTCC	TTACTTTTTT	AATTGATGTT	AGAAGATTCA	GATACCCAGA	GAGACCAATT	960
ATATATTACT	CTGTCTGTGA	CAGCATTGTA	TCTCTTATGT	ACTTCATTGG	ATTTTGTCTG	1020
GGCGATAGCA	CAGCCTGCAA	TAAGGCAGAT	GAGAAGCTAG	AACCTGGTGA	CAGTGTGTCT	1080
CTAGGCTCTC	AAAATAAGGC	TTGCACCGTT	TTTGTCTATG	TTTTCACATG		1140
GCTGGCAGTG	TGTGTGGGTT	GATTCTTACC	ATTACTTGGT	TCTTAGCTGC	AGGAAGAAAA	1200
TGGAGTTTGT	AAGCCATCGA	GCAAAAAGCA	GTGTGTTTTC	ATGCTGTTCG	ATGGGGAACA	1260
CCAGGTTTCC	TGACTGTTAT	GCTTCTTGCT	CTGAACAAAG	TTGAAGGAGA	CAACATTAGT	1320
GGAGTTTGGT	TTGTTGGGCT	TTATGACCTG	GATGCTTCTC	GCTACTTTGT	ACTCTTGCCA	1380
CTGTGCTTTT	GTGTGTTTGT	TGGGCTCTCT	CTTCTTTTAG	CTGGCATTAT	TTCTTAAAT	1440
CATGTTTCAG	AGGTATACAA	ACATGATGGC	CGGAACCAAG	AAAACTAAA	GAATTTTATG	1500
ATTGGAATTG	GAGTCTTCAG	CGGCTTGTAT	CTTGTGCCAT	TAGTGACACT	TCTCGGATGT	1560
TACGCTATAG	AGCAAGTGAA	CAGGATTACC	TGGGAGATAA	CTTGGGTCTC	TGATCATTTG	1620
CGTCAATACC	ATATCCCATG	TCCTTATCAG	GCAAAAAGCA	AAGCTCGACG	AGAATTGGCT	1680
TTATTTATGA	TAAATACCTT	GATGACATTA	ATTGTTGGCA	TCTCTGTGTG	CTTCTGGGTT	1740
GGAAGCAAAA	AGACATGCAC	AGAATGGGCT	GGGTTTTTTA	AACGAAATCG	CAAGAGAGAT	1800
CCAAATCAGTG	AAAGTCGAAG	AGTACTACAG	GAATCATGTG	AGTTTTTCTT	AAGGCACAAT	1860
TCTAAAGTTA	AACACAAAAA	GAAGCACTAT	AAACCAAGTT	CACACAAGCT	GAAGGTCATT	1920
TCCAATCCCA	TGGGAACGAG	CACAGGAGCT	ACAGCAATTC	ATGGCACTTC	TGCAGTAGCA	1980
ATTACTAGCC	ATGATTACCT	AGGACAAGAA	ACTTTGACAG	AAATCCAAAC	CTCACCAGAA	2040
ACATCAATGA	GAGAGGTGAA	AGCGGACGGA	GCTAGCACCC	CCAGGTAAAG	AGAACAGGAC	2100
TGTGGTGAAC	TGCGCTCGCC	AGCAGCATCC	ATCTCCAGAC	TCTCTGGGGA	ACAGTTCGAC	2160
GGGAAGGGCC	AGGCAGGCG	TGATATCTGAA	AGTGCGCGGA	GTGAAGGAAG	GATTAGTCCA	2220
AAGAGTGATA	TTACTGACAC	TGGCCTGGCA	CAGAGCAACA	AATTGACAGT	CCCCAGTTCT	2280
TCGAACCAAA	GCAGCCTCAA	AGGTTCACCA	TCTCTGCTTG	TTCAACCCAGT	TTCAGGAGTG	2340
AGAAAAGAGC	AGGGAGGTGG	TTGTCTTCA	GATACTTGAA	GAACATTTTC	TCTCGTTACT	2400
CAGAAGCAAA	TTTGTGTTAC	ACTGGAAGTG	ACCTATGCAC	TGTTTGTGAA	GAATCACTGT	2460
TACGTTCTTC	TTTTCACATT	AAAGTTGCAT	TGCCTACTGT	TATATCTGAA	AAAATAGAGT	2520
TCAGAAATAA	TATGACTCAT	TTACACAAAA	GGTTAATGAC	AACAATATAC	CTGAAAAACAG	2580
AAATGTGCAG	GTTAATAATA	TTTTTTTAA	AGTGTGGGAG	GACAGAGTTA	GAGGAATCTT	2640
CCTTTCTTAT	TTATGAAGAT	TCTACTCTTG	GTAAGAGTAT	TTTAAGATGT	ACTATGCTAT	2700
TTTACCTTTT	TGATATAAAA	TCAAGATATT	TCTTTGCTGA	AGTATTTAAA	TCTTATCTCT	2760
GTATCTTTT	ATACATATTT	GAAAATAAGC	TTATATGTAT	TTGAACCTTT	TTGAATCCTT	2820
ATTCAAGTAT	TTTTATCATG	CTATTGTGAT	ATTTTAGCAC	TTTGGTAGCT	TTTACACTGA	2880
ATTTCTAAGA	AAATGTGAAA	ATAGTCTTCT	TTTATACTGT	AAAAAAGAT	ATACCAAAAA	2940
GTCTTATAAT	AGGAATTTAA	CTTTAAAAAC	CCACTTATTG	ATACCTTACC	ATCTAAAAATG	3000
TGTGATTTT	ATAGTCTCGT	TTTAGGAATT	TCACAGATCT	AAATATGTGA	ACTGAAATAA	3060
GGTGCTTACT	CAAGAGAGTG	CCACTATTGA	TTGTATTATG	CTGCTCACTG	ATCTTCTCTG	3120
ATATTTAAAA	TAAATGTGCC	TAAAGGGTTA	GTAGACAAAA	TGTTAGTCTT	TTGTATATTA	3180
GGCAAGTGCC	AATTGACTTC	CTTTTAAAA	TGTTTATGTA	CCACCCATTG	ATTGTATTAT	3240
AACCACTTAC	AGTTGCTTAT	ATTTTGTGTT	TTAACCTTTG	TTTCTTAAAC	TTTAGAATAT	3300
TACATTTTGT	ATTATACAGT	ACCTTCTCTA	GACATTTTGT	AG		

Seq ID NO: 651 Protein sequence  
Protein Accession #: NP\_003497.1

1	11	21	31	41	51

MEMFTPLLTTC IFPLPLRGHS LFTCEPITVP RCMKMYNMT FFFNLMGHYD QSIAAEMEH 60  
 FLPLANLECS PNIEPLFCKA FVPTCIEQIH VVPPCRKLC EKVSDCKKLI DTFGIRWPEE 120  
 LBCDRLQYCD ETVPTVTFDPH TEPLGPQKKT EQVQRDIGFW CPRHLKTSBG QGYKFLGIDQ 180  
 CAPPCPNMYF KSDLEFPKKS FIGTVSIFCL CATLFTFLTF LIDVRRFPYP ERPIIYYSVC 240  
 YSIVSLMYFI GFLLDSTAC NKADEKLELG DTVVLGSSQNK ACTVLFMLLY PFTMAGTVWV 300  
 VILTTTWFLA AGRKWSCBAI EQKAVNPHAV AWGTPGFLTV MLLALNKVEG DNISGVCFVG 360  
 LYDLIDASYF VLLPLCLCVF VGLSLLLAGI ISLNHVRQVI QHNGRNQEKL KKFMRIGVFP 420  
 SGLYLVLVLT LLGCVYVEQV NRITWEITWV SDHCROYHIP CPYQAKAKAR PELALFMKY 480  
 LMTLIVGISA VFWVGSKKT TEWAGFPKRN RKRDPISESR RVLQESCEFF LKHSKVKHK 540  
 KKHYKPSHHK LKVISKSMGT STGATANHGT SAVAITSHDY LGQETLTLEIQ TSPETSMREV 600  
 KADGASTPRL REQDCGEPAS PAASISRLSG EQVDGKGQAG SVSESARSEG RISPKSDITD 660  
 TGLAQSNMLQ VPSSEPPSSL KGSTSLLVHP VSGVRKEQGG GCHSDT

Seq ID NO: 652 DNA sequence  
 Nucleic Acid Accession #: NM\_014791.1  
 Coding sequence: 171..2126

1 11 21 31 41 51  
 TTGGCGGGCG GAAGCGGCCA CAACCCGGCG ATCGAAAAGA TTCTTAGGAA CGCCGTACCA 60  
 GCCCGGTCTC TCAGGACAGC AGGCCCTGT CCTTCTGTG GCGCGCGCTC AGCCGTGCCC 120  
 TCCGCCCTC AGGTCTCTTT TCTAATCCA AATAAACTTG CAAGAGGACT ATGAAAGATT 180  
 ATGATGAAC TCTCAATAT TATGAATTAC ATGAAACTTG TGGACAGST GCCTTTGCAA 240  
 AGGTCAAACT TGCTTGCAT ATCCTTACTG GAGAGATGGT AGCTATAAAA ATCATGGATA 300  
 AAAACACACT AGGGAGTGT TTGCCCGGGA TCAAAACGGA GATTGAGGCC TTGAAGAACC 360  
 TGAGACATCA GCATATATGT CAACCTTACC ATGTGCTAGA GACAGCCCAAC AAAATATTCA 420  
 TGGTCTCTGA GTACTGCCCT GGAGGAGAGC TGTTTGACTA TATAATTTC CAGGATCGCC 480  
 TGTCAAGA GAAGACCCGG GTTGTCTTCC GTCAGATAGT ATCTGCTGTT GCTTATGTGC 540  
 ACAGCCAGGG CTATGCTCAC AGGGACCTCA AGCCAGAAAA TTGCTGTTT GATGAATATC 600  
 ATAAATATAA GCTGATTGAT TTTGCTCTCT GTGCAAAACC CAAGGGTAAC AAGGATTACC 660  
 ATCTACAGAC ATGCTGTGGG AGTCTGGCTT ATGCAGCACC TGAGTTAATA CAAGGCAAAAT 720  
 CATATCTTGG ATCAGAGGCA GATGTTGGA GCATGGGCAT ACTGTTATAT GTTCTTATGT 780  
 GTGGATTCTT ACCATTGTAT GATGATAATG TAATGGCTTT ATACAAGAAG ATTATGAGAG 840  
 GAAATATAGA TGTTCCTCAG TGGCTCTCTC CCAAGTAGCAT TCTGCTTCTT CAACAAATGC 900  
 TGCAGGTGGA CCCAAGAAAG CGGATTCTTA TGAATAATCT ATTGAACCAT CCCTGGATCA 960  
 TGCAAGATTA CAACATCTCT GTTGAAGTGC AAAGCAAGAA TCCTTTTATT CACCTCGATG 1020  
 ATGATTGCGT AACAGAACTT TCTGTACATC ACAGAAACAA CAGGCAACA ATGAGGATT 1080  
 TAATTTTCACT GTGGCAGTAT GATCACCTCA CGGCTACCTA TCTTCTGCTT CTAGCCAAGA 1140  
 AGGCTCGGG AAAACCACTT CGTTTAAGGC TTTCTTCTTT CTCTGTGGA CAAGCCAGTG 1200  
 CTACCCCATC CACAGACATC AAGTCAATA ATTGGAGTCT GGAAGATGTG ACCGCAAGTG 1260  
 ATAAAAATAA TGTGCGGGA TTAATAGACT ATGATTGGTG TGAAGATGAT TTATCAACAG 1320  
 GTGCTGTCTC TCCCGAACA TCACAGTTTA CCAAGTACTG GACAGAATCA AATGGGGTGG 1380  
 AATCTAAATC ATTAATCCA GCCTTATGCA GAACACCTGC AAATAAATA AAGAACAAAG 1440  
 AAAATGTATA TACTCTTAAG TCTGCTGTAA AGAATGAAGA GTACTTTATG TTCTCTGAGC 1500  
 CAAAGACTCC AGTTAATAAG AACAGCATA AGAGAGAAAT ACTCACTACG CCAAAATGTT 1560  
 ACACATACAC CACTAAAGCT AGAAACAGT GCTGAAAGA AACTCCAATT AAAATACCAAG 1620  
 TAAATTCAC AGGAACAGAC AAGTTAATGA CAGGTGTCTT TAGCCCTGAG AGGCGGTGCC 1680  
 GCTCAGTGA GTTGGATCTT AACCAAGCAC ATATGGAGGA GACTCCAAAA AGAAGGGGAG 1740  
 CCAAAAGTGT TGGAGCCTT GAAAGGGGAT TGGATAAGST TATCACTGTG CTCACCAGGA 1800  
 GCAAAAGGAA GGGTCTGCCC AGAGACGGGC CCAGAGACT AAAGCTTAC TATAATGTGA 1860  
 CTACAACTAG ATTAGTGAAT CCAGATCAAC TGTGAATGA AATAATGTCT ATTCTTCCAA 1920  
 AGAAGCATGT TGACTTTGTA CAAAAGGGTT ATCACTGAA GTGTCAACA CAGTCAGATT 1980  
 TTGGGAAGT GCAATGCAA TTTGAATTAG AAGTGTGCCA GCTTCAAAA CCGATGTGG 2040  
 TGGGTATCAG GAGGAGCGG CTTAAGGCG ATGCTGCGT TTAACAAGA TTAGTGAAG 2100  
 ACATCCTATC TAGCTGCAAG GTATAATTGA TGGATTCTTC CATCTGCCG GATGAGTGTG 2160  
 GGTGTATAC AGCCTACATA AAGACTGTTA TGATGCTTT GATTTTAAAG TTCATTGGAA 2220  
 CTACCAACTT GTTCTAAG AGCTATCTTA AGACCAATAT CTCTTGTGT TTAACAAAA 2280  
 GATATTATT TTGTATGATA TCTAAATCAA GCCCATCTGT CATATGTGA CTGCTTTT 2340  
 TAATCATGTG GTTTTGTATA TTAATAATTG TTGACTTCT TAGATTCACT TCCATATGTG 2400  
 AATGTAAGCT CTTAACTATG TCTCTTTGTA ATGTGTAATT TCTTCTGAA ATAAACCAT 2460  
 TTGTGAATAT

Seq ID NO: 653 Protein sequence  
 Protein Accession #: NP\_055606.1

1 11 21 31 41 51  
 MKDYDELLKY YELHETIGTG GFAKVKLACH ILTGEMVAIK IMDKNLTGSD LPRIKTEIEA 60  
 LKNLRHOHIC QLYHVLETAN KIPMVLEYCP GGELFDYIIS QDRLESEETR VVFRQIVSAV 120  
 AYVHSQGYAH RDLKPENLLF DEYHKLKID FGLCAKPKGN KDYLQTCOG SLAYAAPELI 180  
 QKSYLGSEA DVWSMGILLY VLMCGFLFPD DDNMALYKK IMRGKYDVPK WLSPPSILL 240  
 QQMLQVDPK RISMKNLLNH PWIMQDYNYP VEWQSKNFFI HLDDDCVTEL SVHHRNRQT 300  
 MEDLISLWQV DELTATYLLL LAKKARGKPV RLRLSSFSFG QASATPPTDI KSNWLSLEDV 360  
 TASDKNTVAG LIDYDWCEDD LSTGAATPRT SQPTKYNTES NGVESKSLTP ALCRTPANKL 420  
 RNKENVTYTPK SAVKNEEYFM FPEPKTPVNK NQHKREILT PNRVYTPSKA RNQCLKETPI 480  
 KIPVNSTGTD KLMTGVISPE RRCSRVELDL NQAHMEETPK RKGAKVGSLS ERGLDKVITV 540  
 LTRSKRKGSA RDGPRRLKHL YNVTTTRLVN PDQLLNEIMS ILPKKHVDFV QKGYTLKQCT 600  
 QSDFGKVTMQ FELEVQQLQK PDVVGIRQR LKGDWVYKR LVEDILSSCK V

Seq ID NO: 654 DNA sequence  
 Nucleic Acid Accession #: NM\_000582  
 Coding sequence: 88..990

1 11 21 31 41 51  
 GCAGAGCACA GCATCGTCGG GACCAGACTC GTCTCAGGCC AGTTGCAGCC TTCTCAGCCA 60  
 AACGCCGACC AAGGAAAAC CACTACCATG AGAATTGCAG TGATTGTCTT TTGCTCCTTA 120

GGCATCACC GTGCCATACC AGTTAAACAG GCTGATTCTG GAAGTTCTGA GGAAAAGCAG 180  
 CTTTACAACA AATACCCAGA TGCTGTGGCC ACATGGCTAA ACCCTGACCC ATCTCAGAAG 240  
 CAGAATCTCC TAGCCCCACA GACCCCTTCCA AGTAAGTCCA ACGAAAGCCA TGACCACATG 300  
 5 GATGATATGG ATGATGAAGA TGATGATGAC CATGTGGACA GCCAGGACTC CATTGACTCG 360  
 AACGACTCTG ATGATGTAGA TGACACTGAT GATTCTCACC AGTCTGATGA GTCTCACCAT 420  
 TCTGATGAAT CGTATGAAT GGTCACTGAT TTTCCACCG AGCTGCCAGC AACCGAAGTT 480  
 TTCACTCCAG TTGTCCCCAC AGTAGACACA TATGATGGCC GAGGTGATAG TGTGGTTTAT 540  
 GGACTGAGGT CAAATCTTAA GAAGTTTCGC AGACCTGACA TCCAGTACCC TGATGCTACA 600  
 10 GACGAGGACA TCACCTCACA CATGGAAAGC GAGGAGTTGA ATGGTGCGTA CAAGGCCATC 660  
 CCGGTGGCCC AGGACCTGAA CGCGCCTTCT GATTGGGACA GCCGTGGGAA GGACAGTTAT 720  
 GAAACGAGTC AGCTGGATGA CCAGAGTGCT GAAACCCACA GCCACAAGCA GTCCAGATTA 780  
 TATAAGCGGA AAGCCAATGA TGAGAGCAAT GAGCATTCCG ATGTGATTGA TAGTCAGGAA 840  
 CTTTCCAAG TCAGCCGTGA ATTCCACAGC CATGAATTC ACAGCCATGA AGATATGCTG 900  
 15 GTTGTAGACC CCAGAAAGTAA GGAAGAAGT AAACACCTGA AATTTCTGAT TTCTCATGAA 960  
 TTAGATAGTG CATCTTCTGA GGTCAATTAA AAGGAGAAAA AATACAATTT CTCACTTTGC 1020  
 ATTTAGTCAA AAGAAAAAAT GCTTTATAGC AAAATGAAAG AGAACATGAA ATGCTTCTTT 1080  
 CTCAGTTTAT TGGTGAATG TGTATCTATT TGAGTCTGGA AATACTAAT GTGTTTGATA 1140  
 ATTAGTTTAG TTTGTGCGTT CATGGAAACT CCCTGTAAAC TAAAAGCTTC AGGGTTATGT 1200  
 20 CTATGTTTAT TCTATAGAG AAATGCAAACT TATCAGTGA TTTTAATATT TGTATTCTC 1260  
 TCAATGATAG AAATTTATGT AGAAGCAAACT AAAATACTTT TACCCACTTA AAAAGAGAT 1320  
 ATAACTTTT ATGCTCACTT AATCTTTTGT TTTTAAAGTT AGTGATATT TGTGTTGATG 1380  
 TATCTTTTGT TGGTGTGAAT AAATCTTTTA TCTTGAATGT AATAAGAAAT TGGTGGTGTG 1440  
 AATTGCTTAT TTGTTTCCC ACGTTGTGCC AGCAATTAAT AAAACATAAC CTTTTTTACT 1500  
 25 GCCTAAAAAA AAAAAAATAA AAAA

Seq ID NO: 655 Protein sequence  
Protein Accession #: NP\_000573

1 11 21 31 41 51  
 30 MRIAVICFCL LGITCAIPVK QADSGSSEK QLYNKYPDAV ATWLNPDPSQ KQNLAPQTL 60  
 PSKSNESHDH MDMDDDDDD DHVDSQDSID SNDSDDVDVT DSHQSDSH HSDESDELVT 120  
 DFFTLDPATE VFTFVPVPTVD TYDGRGDSVV YGLRSKSKKF RRPDIQYFDA TDEDITSHME 180  
 35 SEELNGAYKA IPVAQDLNAP SDWDSRGKDS YETSQLDDQS AETHSEKQSR LYKRKANDES 240  
 NEHSDVIDSQ ELKSVSREFF SHEFHSHEDM LVVDPKSKEE DKELKFRISH ELDSASSEVN

Seq ID NO: 656 DNA sequence  
Nucleic Acid Accession #: NM\_003108.1  
Coding sequence: 76..1401

1 11 21 31 41 51  
 45 GGGGTGGGAG GGGGAGGGGG ACCTCCGCAC GAGACCCAGC GGCCCGGGTT GGAGCGTCCA 60  
 GCCTCGCAAC GGATCATGGT GCAGCAGGCG GAGAGCTTGG AAGCGGAGAG CAACCTGCCC 120  
 CGGGAGGGCG TCAGACACGGA GGAGGGGGAA TTCTAGGCTT GCAGCCCGGT GCCCTGGAC 180  
 GAGAGCGACC CAGACTGGTG CAAGAAGGCG TCGGGCCACA TCAAGCGGCC GATGAACCGG 240  
 50 TTCAATGATG GGTCCAAGAT CGAACGAGG AAGATCATGG AGCAGTCTCC GGACATGCAC 300  
 AACGCCGAGA TCTCCAAGAG GCTGGGCAAG CGCTGGAAAA TGCTGAAGGA CAGCGAGAAG 360  
 ATCCCGTTCA TCGGGGAGGC GGAGCGGCTG CGGCTCAAGC ACATGGCCGA CTACCCCGAC 420  
 TACAAGTACC GGCCCGGAA AAGGCCCAA ATGGACCCCT CGGCCAAGCC CAGCGCCAGC 480  
 CAGAGCCAG AGAAGAGCGC GGCCCGCGGC GCGCGCGGGA GCGCGGGCGG AGCGCGGGG 540  
 GGTGCCAAGA CTTCCAAGGG CTCCAGCAAG AAATGCGGCA AGCTCAAGGC CCGCGCGGC 600  
 55 GCGGGCGCCA AGCGGGCGC GGGCAAGGCG GCGCAGTCCG GGGACTACGG GGGCGCGGGC 660  
 GAGGACTACG TGCTGGGCGC CCGTGGCGTG AGCGGCTCGG GCGCGCGCGG CGCGGCGAAG 720  
 ACGGTCAAGT CGGTGTTTCT GGATGAGGAC GAGGAGGAG ACGAGCAGCA CGACGAGCTG 780  
 CAGCTCGAGA TCAAAACAGGA GCCGAGCAG GAGGAGGAG AACACCGCA CCAGCAGCTC 840  
 CTGCGAGCGC CGGGGAGCAG GCCGTGCGAG CTGCTGAGAC GCTACAAAGT CGCCAAAGTG 900  
 60 CCGCGCAGC CTACGCTGAG CAGCTCGGCG GAGTCCCCCG AGGGAGCGAG CCTCTACGAC 960  
 GAGGTGCGCG CGGCGCGGAC CTGCGCGGCC GGGGGCGGCA GCGGCTCTTA CTACAGCTTC 1020  
 AAGAACATCA CCAAGCAGCA CCGCGCGCGC CTCGCGCAGC CCGCGCTGTC GCCCGGCTCC 1080  
 TCGCGCTCGG TGTCCACCTC CTGCTCCAGC AGCAGCGGCA GCAGCAGCGG CAGCAGCGGC 1140  
 GAGGACGCG AGCAGCTGAT GTTCGACCTG AGCTTGAATT TCTCTCAAAG CGCGCACAGC 1200  
 65 GCCAGCGAGC AGCAGCTGGG GGGCGCGCGC GCGCGCGGGA ACCTGTCCCT GTCGCTGGTG 1260  
 GATAAGGATT TGGATTCTGT CAGCGAGGCG AGCCTGGGCT CCACTTCCA GTTCCCGGAC 1320  
 TACTGCACGC CGGAGCTGAG CGAGATGATC CGGCGGAGT GGTGAGGCG GAACCTCTCC 1380  
 GACCTGGTGT TCACATATTG AAAGGCGGCC GCTGCTCGCT CTTTCTCTCG GAGGGTGCAG 1440  
 AGCTGGGTTT CTTGGGAGGA AGTTGTAGTG GTGATGATGA TGATGATGAT AATGATGATG 1500  
 70 ATGATGGTGG TGTGTAGTGT GCGGTTGGTA GGGTGGAGGG GAGAGAAGAA GATGCTGATG 1560  
 ATATTGATAA GATGTCGTGA CGCAAGAAA TTGGAAAAA TGATGAAAT TTTGGTGGAG 1620  
 TTAAGTGAA ATGAGTAGTT TTTAAACATT TTCTCTGTC TTTTCTGTC CCCCCTCCCT 1680  
 TCCCTTATCG TGTCTCAAGG TAGTTGCATA CCTAGTCTGG AGTTGTGATT ATTTCCCAA 1740  
 AAAATGTGTT TTTGTAATTA CTATTTCTTT TTCTGAAAT TCGTGATTGC AACAAAGGCA 1800  
 75 GAGGGGCGCG CGCGGCGGAG GGGAGGTAGG ACCCGCTCCG GAAGCGCTG TTTGAAGCTT 1860  
 GTGCTCTTT GAAGTCTGGA AGACGCTGCG AGAGGACCTT TTTGGCAGCA CAACGTGTAC 1920  
 TCTAGGGAGT TGGTGGAGAT ATTTTCTTTT CTTAAGAGAA CTTAAAGAAC TGGTGAATTT 1980  
 TTTTAAACAA AAAAAGGG

Seq ID NO: 657 Protein sequence  
Protein Accession #: NP\_003099.1

1 11 21 31 41 51  
 85 MVQAESLEA ESNLPREALD TEEGEPMACS FVALDESDPD WCKTAGSHIK RPNMNPVWS 60  
 KIERRIMBQ SPDMHNAETS KRLGKRWKML KDSEKIPFIR EAERLRLKHM ADYDYKYRP 120  
 RKFKMDPSA KPSASQSPFK SAAGGGGSGA GGGAGGAKTS KSSSKKCGKL KAPAAAGAKA 180  
 GAGRAAGSD YGAGDDYVL GSLRVSGSGG GGAGKTIVKCV FLDEDDDDDD DDELQLQIK 240

QEPDEDEREP PHQQLQPPG QQPSQLLRRY NVAKVPASPT LSSSAESPEG ASLYDEVVRAG 300  
 ATSGAGGGR LYYSPKNITK QHPPLAQA LSPASSRSVS TSSSSSSGSS SGSSGEDADD 360  
 LMFDSLNLFS QSAHSASEQQ LGGGAAGNL SLSLVDKDLDF SFSEGLGSH FEPPDYCTPE 420  
 LSEMIAGDWL EANPSDLVFT Y

Seq ID NO: 658 DNA sequence  
 Nucleic Acid Accession #: NM\_001719  
 Coding sequence: 123..1418

1 11 21 31 41 51  
 | | | | |  
 GGGCGCAGCG GGGCCCCGTCT GCAGCAAGTG ACCGACGGCC GGAACGGCCG CTGCCCCCT 60  
 CTGCCACCTG GGGCGGTGCG GGCCCGGAGC CGGAGGCCCG GTTAGCGCGT AGAGCCGGCG 120  
 CGATGCACGT GCGCTCACTG CGAGCTGCGG CGCCGCACAG CTTCGTGGCG CTCTGGGCAC 180  
 CCCTGTCTCT GCTGCGCTCC GCGCTGGCGG ACTTCAGCCT GGACCAACGAG GTGCACTCGA 240  
 GCTTCATCCA CGGGCGCTC CGCAGCCAGG AGCGGCGGGA GATGCAGCGC GAGATCCTCT 300  
 CCATTTTGGG CTTGCCCCAC GCGCCGCGCC GGCACCTCCA GGGCAAGCAC AACTCGGCAC 360  
 CCATGTTTAT GCTGGACCTG TACAACGCCA TGGCGGTGGA GGAGGCGCGG GGGCCCGCGG 420  
 GCCAGGCTT CTCTACCCG TACAAGGCCG TCTTCAGTAC CCAGGGCCCC CCTCTGGCCA 480  
 GCCTGCAAGA TAGCCATTTC CTCACCGAGC CGACATGGT CATGAGCTTC GTCAACCTCG 540  
 TGGAAACATGA CAGGAATTTC TTCCACCCAC GCTACCACCA TCGAGAGTTC CGGTTTGATC 600  
 TTTCCAGAT CCCAGAGGGG GAAGCTGTCA CGGCAGCGGA ATTCCGATC TACAAGGACT 660  
 ACATCCGGGA AGCTTTCGAC AATGAGAAGT TCCGATCAG CTTTATCAG GTGCTCCAGG 720  
 AGCACTTGGG CAGGGAATCG GATCTCTTCC TGCTCGACAG CCGTACCCCT TGGGCTCGG 780  
 AGGAGGGCTG GCTGTGTTT GACATCAGAG CCACAGCAA CCACTGGGTG GTCAATCCGC 840  
 GGCACAACT CAGGCTGAGC CTCTCGGTGG AGACGTGGA TGGGCGAGG ATCAACCCCA 900  
 AGTTGGCGGG CTGATTTGGG CGGCACGGGC CCCAGAACAA GCAGCCCTTC ATGGTGGCTT 960  
 TCTTCAAGGC CACGAGGTC CACTTCCGCA GCATCCGTC CACGGGAGC AAACAGCGCA 1020  
 GCCAGAACCG CTCAGAGCG CCCAAGAAC AGGAAGCCCT GCGATGGCC AACGTGGCAG 1080  
 AGAACAGCAG CAGCAGCAG AGGCAGGCT GTAAAGAACGA CAGCTGTAT GTCACTTCC 1140  
 GAGACCTGGG CTGGCAGAGC TGGATCATCG CGCTGAAAG CTACGCGCC TACTACTGTG 1200  
 AGGGGAGTG TGCTTCCCT CTGAATCCT ACATGAACGC CACCAACCCAC GCCATCGTGC 1260  
 AGACGCTGGT CCATTTTATC AACCCGAAA CGGTGCCCAA GCCCTGCTGT GCGCCACGC 1320  
 AGCTCAATGC CATCTCCGTC CTCTACTTCG ATGACAGCTC CAACTCATC CTGAAGAAAT 1380  
 ACAGAAACAT GGTGTTCGGG GCTCTGGCT GCCACTAGCT CCTCCGAGAA TTCAGACCTT 1440  
 TTGGGGCCAA GTTTTCTGG ATCCTCCATT GCTGCGCTTG GCCAGGAAC AGCAGACCAA 1500  
 CTGCTTTTGG TGAGACTTCT CCTCCCTAT CCGCAACTTT AAAGGTGTGA GAGTATTAGG 1560  
 AAACATGAGC AGCATATGGC TTTTGATCAG TTTTTCAGTG GCAGCATCCA ATGAACAGA 1620  
 TCCTACAAGC TGTGCAGGCA AAACCTAGCA GGAAGAAAA ACAACGCATA AAGAAAAATG 1680  
 GCGCGGCCAG GTCACTGGCT GGAAGTCTC AGCCATGCAC GGACTCGTTT CCAGAGGTAA 1740  
 TTATGAGCGC CTACAGCAGC GGCCACCCAG CGGTGGGAGG AAGGGGCGGT GGCAAGGGGT 1800  
 GGGCACATTG GTGCTGTGC GAAAGGAAAA TTGACCCGGA AGTTCTGTGA ATAAATGTCA 1860  
 CAATAAACCG AATGAATG

Seq ID NO: 659 Protein sequence  
 Protein Accession #: NP\_001710

1 11 21 31 41 51  
 | | | | |  
 MHVRSRAAA PHSFVALWAP LFLRLSALAD FSLDNEVHSS FIHRLRLRQRE REMQREILS 60  
 ILGLPERPRP HLQGGKNSAP MFMLDLYNAM AVEEGGGPGG QGFSYPYKAV PSTQGPPLAS 120  
 LQDSHFLTDA DMVMSFVNLV EHDKEFFHPR YHREFRFDL SKIPEGBAVT AAEPRIVKDY 180  
 IRERFDNFT RISVYQVLQE HLGRESDLFL LDSRTLWASE EGWLVFDITA TSNHWVWNP 240  
 HNLGLQLSVE TLDGGSNPK LAGLGRHGP QNKQPFMVAF FKATVEHFRS IRSTGEKORS 300  
 QNRSTPKNQ BALRMANVAE NSSSDQRQAC KKHLYVSVFR DLGNQDWIIA PEGYAAYYCE 360  
 GECAPPLNSY MNATHAIVQ TLVHFINPET VFKPCCAPTQ LNAISVLYFD DSSNVILKRY 420  
 RNMVVRACGC H

Seq ID NO: 660 DNA sequence  
 Nucleic Acid Accession #: E05 sequence  
 Coding sequence: 211..1895

1 11 21 31 41 51  
 | | | | |  
 GGATCTGAGG GGGCGCCAGT CACTTCTCTC AGTTCTCTGT GCTGGGCGGG AGGAGCGGAT 60  
 GGGGCTTGGG AGGCAGCCTG CTCTCCAGTC CCTATCCACC CACAGGTTTT TTGGGTGGGA 120  
 GAGGAATTAT CTGATAAAAT TCCTGGGTTA ATATTTTIAA AAACGGAGAG TTTTAAAAA 180  
 TGATTTTTTT CCCTCGAAAA TGACCTTTTT ATGCTTCGAA GCAGTTTGTC AACCAAGATA 240  
 GTGCTTTTTT TTTTCTCTTC TTTTCTACG ATAAATGAAA GCATTTCTTC AAGAAAAAGG 300  
 CACAGGTTCC TTGAACAGCT GGATTCGAT GGCACCATTA CTATAGAGGA GCAGATTGTC 360  
 CTTGTCTGA AAGCGAAAGT ACAATGTGAA CTCAACATCA CAGCTCAACT CCAGGAGGGA 420  
 GAAGGTAATT GTTCCCTGA ATGGGATGGA CTCAATTTGT GGGCCAGAGG AACAGTGGG 480  
 AAAAATACGG CTGTTCCATG CCTCCTTAT ATTTATGACT TCAACCATAA AGGAGTTGCT 540  
 TTCCGACACT GTAACCCCAA TGAACATGG GATTTTATGC ACAGCTTAAA TAAACATGG 600  
 GCCAATTATT CAGACTGCC TCGCTTCTG CAGCCAGATA TCAGCATAGG AAAGCAAGAA 660  
 TTCTTGAAC GCTCTATGT AATGTATACC GTTGGCTACT CCATCTCTT TGGTTCCCTG 720  
 GCTGTGGCTA TTCTCATCAT TGGTACTTTC AGACGATTGC ATTGCACTAG GAATATATC 780  
 CACATGCACT TATTTGTGCT TTTCATGCTG AGAGCTACAA GCATCTTGT CAAAGCAGA 840  
 GTAGTCCATG CTCACATAGG AGTAAAGGAG CTGGAGTCCC TAATAATGCA GGATGACCCA 900  
 CAAAATCCA TTGAGGCAAC TTCTGTGGAC AAATCACAAT ATATCGGGGT CAAGATTGCT 960  
 GTTGTGATGT TATTTACTT CTGCTGTACA AATTATTAT GGATCCTGGT GGAAGGTCTC 1020  
 TACCTGCATA ATCTCATCTT TGTGGCTTTC TTTTCGACA CCAATACCT GTGGGGCTTC 1080  
 ATCTTGATAG GCTGGGGTTT TCCAGCAGCA TTGTGTGCG CATGGGCTGT GGCACGAGCA 1140  
 ACTCTGGCTG ATGCGAGGTG CTGGGAACCT AGTGTCTGAG ACATCAAGTG GATTATCAA 1200  
 GCACCGATCT TAGCAGCTGT TGGGCTGAAT TTTATCTGT TTCTGAATC GGTTAGAGTT 1260  
 CTAGCTACCA AAATCTGGGA GACCAATGCA GTTGGCATG ACACAAGGAA GCAATACAGG 1320  
 AAATGSCCA AATGCAGACT GGTCTGGTC CTAGTCTTTG GAGTGATTA CATCTGTTC 1380

OTATGCCCTGC CTCACTCCTT CACTGGGCTC GGGTGGGAGA TCCGCATGCA CTGTGAGCTC 1440  
 TTCTTCAACT CCTTTCAGGG TTCTTTGTG TCTATCATCT ACTGCTACTG CAATGGAGAG 1500  
 GTTCAGGCAG AGGTGAAGAA GATGTGGAGT CGGTGGAATC TCTCCGTGGA CTGGAAGAGG 1560  
 5 ACACCGCCAT GTGGCAGCGG CAGATGCGGC TCAGTGCTCA CCACCGTGAC GCACAGCACC 1620  
 AGCAGCCAGT GCAGGTGGC GCCCAGCACA CGCATGGTGC TTATCTCTGG CAAAGCTGCC 1680  
 AAGATCCCA GCAGACAGCC TGACAGCCAC ATCACTTTAC CTGGCTATGT CTGGAGTAAC 1740  
 TCAGAGCAGG ACTGCTTACC ACATCTTTC CACGAGGAGA CCAAGGAAGA TAGTGGGAGG 1800  
 CAGGGAGATG ATATTCTAAT GGAGAAGCCT TCCAGGCCTA TGAATCTAA CCCAGACACT 1860  
 10 GAAGGATGCC AAGGAGAAAC TGAGGATGTT CTCTGA

Seq ID NO: 661 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 15 MLRSSLSSTSI VLFLPSSFSF INESISSRKR HRPLEQLDSG GTITIEBQIV LVLKAKVQCE 60  
 LNITAQLEQG EGNCFPEWDG LICWPRGTVG KISAVPCPPY IYDFNHKGVA FRHCPNPGTW 120  
 DPMHSLNKTW ANYSDCLRFQ QDISIGQEF FERLYVMYTV GYSISFGSLA VAILIIGYPR 180  
 20 RRLHCTRNYYI HMLFVSPMLR RATSIFVKDR VVHAHIGVKE LESLIMQDDP QNSIEATSVK 240  
 KSQYIGCKIA VUMFIYPLAT NYWILVEGL YLHNLIFVAF FSDTKYLWGF ILIGWGFPA 300  
 FVAAMAVARA TLADARCWEL SAGDIKWYIQ APILAAIGLN FILPLNTVRV LATKIWETNA 360  
 VGHDRKQYR KLAKSTLLVLV LVFGVHYIVF VCLPHSFTGL GWEIRMHCEL PFNSFQGFV 420  
 SIYCYCNGE VQAEVKMWS RWNLSVDWKR TPPCGSRRC SVLTTVTHTS SSQSQAAS 480  
 25 RMVLISGKAA KIASRQPDHS ITPGYVWSN SEQDCLPHSF HEETKEDSGR QGDDILMEKP 540  
 SRPMESNPDT EGCQGETEDV L

Seq ID NO: 662 DNA sequence  
 Nucleic Acid Accession #: NM\_005048  
 Coding sequence: 143..1795

1 11 21 31 41 51  
 30 GGCCGGTGGC CCGGGCCCGA CCACCCAGC TGGCGTCTGT TACTGGCCAC AAGTTTGCTC 60  
 TGGGCCAGCC AAGTTGGCAA CTGGGAAGCT TCTCCGGGC TCTGGAGGAG GGTCCCTGCT 120  
 35 TCTTCTTACA GCGGTTCCGG GCATGGCCGG GCTGGGGCGC TCGTCCACG TCTGGGGTGT 180  
 GCTAATGCTC GGCAGCTGCC TCCTGGCCAG AGCCAGCTG GATTCTGATG GCACCATAC 240  
 TATAGAGGAG CAGATTGTCC TTGTGCTGAA AGCGAAAGTA CAATGTGAAC TCAACATCAC 300  
 AGCTCAACTC CAGGAGGGAG AAGGTAATG TTTCCTGAA TGGATGGAC TCATTGTGTG 360  
 40 CCCCAGAGGA ACAGTGGGGA AATATCGGC TGTTCATGC CCTCCTTATA TTTATGACTT 420  
 CAACCTATAA GGAGTTGCTT TCCGACACTG AACCCCAAT GGAACATGGG ATTTTATGCA 480  
 CAGCTATAAT AAAACATGGG CCAATTATTC AGACTGCCCT CGCTTCTGTC AGCCAGATAT 540  
 CAGCATAGGA AAGCAAGAA TCTTTGAACG CCTCTATGTA ATGTATACCG TTGGCTACTC 600  
 CATCTCTTTT GGTTCTCTGG CTGTGGCTAT TCTCATCAIT GGTACTTCA GAGGATTGCA 660  
 45 TTGCACTAGG AACTATATCC ACATGCACCT ATTTGTGTCT TTCATGCTGA GAGCTACAAG 720  
 CATCTTTGTC AAAGACAGAG TAGTCCATGC TCACATAGGA GTAAAGGAGC TGGAGTCCCT 780  
 AATAATGCAG GATGACCCAC AAAATCCAT TGAGGCACT TCTGTGGACA AATCACAATA 840  
 TATCGGGTGC AAGATTGCTT TTGTGATGTT TATTACTTTC CTGGCTACAA ATTATTATTG 900  
 GATCCTGGTG GAAGTCTCTT ACCTGCATAA TCTCATCTTT GTGGCTTTCT TTTCCGACAC 960  
 50 CAAATACCTG TGGGCTTCA TCTTGATAGG CTGGGGTGT CCAGCAGCAT TTGTGCGAGC 1020  
 ATGGGCTGTG GCACAGACAA CTCTGGCTGA TGGAGGTGC TGGGAACCTA GTGCTGAGAG 1080  
 CATCAAGTGG ATTTATCAAG CACGATCTT AGCAGCTATT GGGCTGAATT TTATTCTGTT 1140  
 TCTGAATACG GTTAGAGTTC TAGCTACCAA AATCTGGGAG ACCAATGCAG TTGGGCATGA 1200  
 CACAAGGAAG CAATACAGGA AACTGGCCAA ATCGACACTG TCTCTGGTCC TAGTCTTTGG 1260  
 55 AGTGCAATAC ATCGTGTTCG TATGCTGCTC TCACTCCTTC ACTGGGCTCG GGTGGGAGAT 1320  
 CCGCATGCAC TGTGAGCTCT TCTTCAACTC CTTTCAGGGT TTCTTTGTGT CTATCATCTA 1380  
 CTGCTACTGC AATGGAGAGG TTCAGGCAGA GGTGAAGAAG ATGTGGAGTC GGTGAATCT 1440  
 CTCGCTGAGC TGGAAAGAGA CACCGCATG TGGCAGCCGC AGATGCGGCT CAGTGTCTAC 1500  
 CACCGTGAGC CACAGCACA CAGCCAGTC ACAGGTGGCG GCCAGCACAC GCATGCTGCT 1560  
 60 TATCTCTGCG AAGCTGCCA AGATCGCCAG CAGACAGCCT GACAGCCACA TCACCTTACC 1620  
 TGGCTATGTC TGGAGTAAT CAGAGCAGGA CTGCTGCCA CACTCTTCC ACGAGGAGAC 1680  
 CAAGGAAGAT AGTGGGAGGC AGGGAGATGA TATTCTAATG GAGAAGCCTT CCAGGCCTAT 1740  
 GGAATCTAAC CCAGACACTG AAGGATGCCA AGGAGAAACT GAGGATGTTT TCTGAATGGA 1800  
 CATTTGTGCG TGACTTTCAT GGGCTGGTCC AATGGCTGGT TGTGTGAGAG GGCTTGGCTG 1860  
 65 ATACTCCTAT GCTTGAGTTC AAAGGCTGAA AATTCAGTTA AGGTGTTACT TAATAATAGT 1920  
 TTTTAGGCTC CATGAATTG CTCTGTAAA TACTAACGAC ATGAAATATG AAGTGTCAAT 1980  
 GGAGTAGTTT ATTACCTCTT ATGGCATCA AGTTTCTCTC TAAATTAATG TATGGTATTT 2040  
 GCTCTGTGAT TGTTCAATTT TTTCTGCTAC TTTTGGGTAG AAAAAAGATT CAATTGCTTG 2100  
 GCTGTAGCTT TCTCTCATAT ATATCACCTT AAATATAATG AAGATCTTTT AGTGTGTATC 2160  
 70 ATTTTCTTT TAGAACTAG TATTCTCTTA TTTCTTACT TAATGTACTT CTATCACTGC 2220  
 ATTTATTTTG CCTGTGCATA GGAGCAATTA GGATCTAAAA AAATATATGG GAAGATAAAA 2280  
 GATCTAAGAA CAGTACTTTC CTGGAATAAT AGTTGGCTGG ACATTGATAA AATAATGCAT 2340  
 TTATAACAAT TACATGTGTT TTTGGGAACA AGGAAATTT CTCAAAAAAG AATATTTCAC 2400  
 ACATCCCTTC TTTTGAATGG CCTCTTTGTG ACCAGCCAGA CCTCAGGTCT TCACCTCTTC 2460  
 75 TCTTTGTAA ACCATGTCTAT GTGGAAAGAT TTCCTCAGT AGTGAGCTTG TGTCTGCAAA 2520  
 TTGATTTGT TGTAAATGTA TTTGATAGC AAATCATGCT GCATCTATAT CTTTCTCTTG 2580  
 TTTGAGCTGT TACTACATTG TACTGTCAT GTGGGATCAA TTAATAATTT GTTTTAAAAA 2640  
 T

Seq ID NO: 663 Protein sequence  
 Protein Accession #: NP\_005039

1 11 21 31 41 51  
 85 MAGLGASLHV WGNWMLGSL LARAQLDSG TITIEBQIVL VLKAKVQCEL NITAQLEQGE 60  
 GNCFFPEMDGL ICWPRGTVGK ISAVPCPPYI YDFNHKGVA FRHCPNPGTW FMHSLNKTWA 120  
 NYSDCLRFQ QDISIGQEF FERLYVMYTV GYSISFGSLA VAILIIGYPR RHLHCTRNYYI 180  
 MHLFVSPMLR RATSIFVKDR VVHAHIGVKE LESLIMQDDP QNSIEATSVK SOYIGCKIAV 240

VMPYIPLATN YYNIVLEGLY LHNLIPIVAFV SDTKYLMGPI LIGWGFPAAF VAANAVARAT 300  
 LADARCWELS AGDIKNIYQA PILAAIGLNP ILFLNTRVRL ATKIWETNAV GHDTRKQYRK 360  
 LAKSTLVLVL VFGVYIYFV CLPHSFTGLG WEIRMHCELF FNSPQGFVVS IYCYCNGEV 420  
 QAEVKIOWSR WNLSDVDMKRT PPGSRRRCQS VLTTVTHSTS SQSQVAASR MVLISGKAAR 480  
 IASRQFDSHI TLEGVVWSNS EQDCLPHSFH EETKEDSGRQ GDDILMEKPS RPFMESNPDE 540  
 GQGETEDVL

Seq ID NO: 664 DNA sequence  
 Nucleic Acid Accession #: NM\_012152  
 Coding sequence: 43..1104

1	11	21	31	41	51	
CTTCTTTAA	TTCTTTCTA	GGATGTTAC	TTCTTCTCA	CAATGAATGA	GTGTCACTAT	60
GACAAGCACA	TGGACTTTTT	TTATAATAGG	AGCAACACTG	ATACTGTGCA	TGACTGGACA	120
GGAAACAAGC	TTGTGATGT	TTTGTGTGT	GGGACGTTT	TCTGCTGTT	TATTTTTTT	180
TCTAATCTC	TGGTCATCGC	GGCAGTGATC	AAAAACAGAA	AATTTCAATT	CCCTTCTAC	240
TACCTGTGG	CTAATTTAGC	TGCTGCCGAT	TTCTTCGCTG	GAATTGCCTA	TGTATTCTCG	300
ATGTTTAA	CAGGCCAGT	TTCAAAAAC	TTGACTGTCA	ACCGCTGGTT	TCTCGTCAG	360
GGGCTTCTG	ACAGTAGCTT	GACTGCTTCC	CTCACCAACT	TGCTGGTTAT	CGCCGTGGAG	420
AGGCACATGT	CAATCATGAG	GATGCGGGTC	CATAGCAACC	TGACCAAAA	GAGGGTGACA	480
CTGCTCAATT	TGCTTGTCTG	GGCCATCGCC	ATTTTATGG	GGGCGGTCCC	CACACTGGGC	540
TGGAATTGCC	TCTGCAACAT	CTCTGCCCTG	TCTTCCCTGG	CCCCCATTTA	CAGCAGGAGT	600
TACCTTGT	TCTGGACAGT	GTCCAACCTC	ATGGCCTTCC	TCATCATGGT	TGTGGTGTAC	660
CTGGGATCT	ACGTGTACGT	CAAGAGGAAA	ACCAACGCTC	TGTCTCGGCA	TACAAGTGGG	720
TCCATCAGCC	GCGGAGGAGT	ACCCATGAAG	CTAATGAAGA	CGGTGATGAC	TGTCTTAGGG	780
GCGTTGTGG	TATGCTGAC	CCCGGCCCTG	GTGGTCTGCG	TCTTGACGG	CCTGAACCTG	840
AGGCAGTGTG	GCGTGCAGCA	TGTGAAAAGG	TGGTCTCTGC	TGCTGGCGCT	GCTCAACTCC	900
TCGTGAACC	CCATCATCTA	CTCCTACAAG	GACGAGGACA	TGTATGGCAC	CATGAAGAAG	960
ATGATCTGCT	GCTTCTCTCA	GGAGAACCCA	GAGAGGCGTC	CCTCTCGCAT	CCCTTCCACA	1020
GTCTTCAGCA	GAGATGACAG	AGGCAGCCAG	TACATAGAGG	ATAGTATTAG	CCAAGTGCA	1080
GTCTGCAATA	AAAGCACTTC	CTAAACTCTG	GATGCTCTC	GGCCACCCA	GGTGATGACT	1140
GTCTTAGG						

Seq ID NO: 665 Protein sequence  
 Protein Accession #: NP\_036284

1	11	21	31	41	51	
MNECHYDKHM	DFPYNRSNTD	TVDDWTGTL	VIVLCVGTFF	CLFIFFSNLS	VIAAVIKNRK	60
FHPFPYLLA	NLAADFFAG	IAYVLFMNT	GPVSKLTVN	RWFLRQLLD	SSLTASLTNL	120
LVIIVERHMS	IMRMVRHNSL	TKKRVTLIL	LWNAIAIFMG	AVPTLGNWCL	QNISACSSLA	180
PYISRYLVF	WTVSNLMAFL	IMVVYLRIY	VYVVKRTNVL	SPHTSGSISR	RRTPMKLMKT	240
VMTVLGAFV	WTPGLVLL	LDGLNCRQCG	VQHVKRWFLL	LALLNSVVP	IYYSKDEDM	300
YGTMKKMIC	FSQENPERRP	SRIPSTVLSR	SDTGSQYIED	SISQGAVCNK	STS	

Seq ID NO: 666 DNA sequence  
 Nucleic Acid Accession #: NM\_002821  
 Coding sequence: 150..3362

1	11	21	31	41	51	
AACTCCGCC	TGCGGAAGCC	TGCGGGTCCG	GCTCCGCTGC	CGGCTGCTGC	TGCGGCGCCC	60
GGCTCCGGT	GGCTCCGCT	CCTGTGCCCG	CCGCGGAGCA	GTCTGCGGCC	CGCCGTGGGC	120
CCTCAGCTCC	TTTTCTGAG	CCCGCCGCA	TGGGAGCTGC	GCGGGGATCC	CCGGCCAGAC	180
CCGCGCGGT	GGCTCTGCTC	AGCGTCTGCT	TGCTGCGGCT	GCTGGGCGGT	ACCCAGACAG	240
CCATTGTCT	CATCAAGCAG	CGGTCTCTCC	AGGATGCACT	GCAGGGGCGC	CGGGCGCTGC	300
TTGCGTGTGA	GTTTGAAGCT	CCGGGCCCGG	TACATGTGTA	CTGGCTGCTC	GATGGGGCCC	360
CTGTCCAGGA	CACGAGGCGG	CGTTTCGCCC	AGGGCAGCAG	CCTGAGCTTT	GCAGCTGTGG	420
ACCAGCTGCA	GGACTCTGCG	ACCTTCCAGT	GTGTGGCTCG	GAATGATGTC	ACTGGAGAAG	480
AAGCCCGCAG	TGCCAAGGCC	TCCTTCAACA	TCAAATGGAT	TGAGGCAGGT	CCTGTGGTCC	540
TGAAGCATCC	AGCTTCGGA	GCTGAGATCC	AGCCACAGAC	CCAGGTGACA	CTTCTGTGCC	600
ACATTGATGG	GCACCTCGG	CCCACTTACC	AATGTTCCG	AGATGGGACC	CCCTTTCTG	660
ATGGTCAGAG	CAACCAACA	GTCAGCAGCA	AGGAGCGGAA	CCTGACGCTC	CGGCCAGCTG	720
GTCTTGAGCA	TAGTGGGCTG	TATTCTGCT	GCGCCACAG	TGCTTTTGGC	CAGGCTTGCA	780
GCAGCCAGAA	CTTACCTTG	AGCATTGTCT	ATGAAAGCTT	TGCCAGGGTG	GTGCTGGCAC	840
CCCGACAGCT	GGTAGTAGCG	AGGTATGAGG	AGGCCATGTT	CCATTGCCAG	TTCTCAGCCC	900
AGCCACCCCG	GAGCTGACAG	TGGCTCTTTG	AGGATGAGAC	TCCCATCACT	AACCGCAGTC	960
CGCCCCCA	CCTCCGAGA	GCCACAGTGT	TTGCCAACGG	GTCTCTGCTG	CTGACCCAGG	1020
TCGCGCCACG	CAATGCAGGG	ATCTACGCT	GCAATTGGCA	GGGGCAGAGG	GGCCACCCA	1080
TCATCTTGA	AGCCCACTT	CACCTAGCAG	AGATTGAAGA	CATGCCGCTA	TTTGAGCCAC	1140
GGGTGTTTAC	AGCTGGCAGC	GAGGAGCGTG	TGACCTGCCT	TCCCCCAAG	GGTCTGCCAG	1200
AGCCACAGGT	GTGGTGGGAG	CACGCGGGAG	TCCGGCTGCC	CACCCATGGC	AGGGTCTACC	1260
AGAAGGGCCA	CGAGCTGGTG	TTGGCCAATA	TTGCTGAAAG	TGATGCTGGT	GTCTACACCT	1320
GCCACGCGCG	CAACCTGGCT	GGTCAGCGGA	GACAGGATGT	CAACATCACT	GTGGCCACTG	1380
TGCCCTCCTG	GCTGAAGAAG	CCCAAGACA	GCCAGCTGGA	GGAGGCAAA	CCGGCTACT	1440
TGGATTGCTT	GACCCAGGCT	ACACCAAAAC	CTACAGTTGT	CTGGTACAGA	AACCAAGATG	1500
TCATCTCAGA	GGACTCAGG	TTGAGGTCT	TCAAGATGCT	GACCTTGGCG	ATCAACAGCG	1560
TGGAGGTGTA	TGATGGGACA	TGGTACCGTT	GTATGAGCAG	CACCCAGGCC	GGCAGCATCG	1620
AGCGCAAGC	CGGTGTCCAA	GTGCTGGAAA	AGCTCAAGTT	CACACCAACA	CCCCAGCCAC	1680
AGCAGTGCAT	GGAGTTTACG	AAGGAGGCCA	CGGTGCCCTG	TTCAGCCACA	GGCCAGAGA	1740
AGCCCACTAT	TAAAGTGGAA	CGGGCAGATG	GGAGCAGCCT	CCCAGAGTGG	GTGACAGACA	1800
ACGCTGGGAC	CTGTGATTTT	GCCCAGGTGA	CTGAGATGA	CGCTGGCAAC	TACACTTGCA	1860
TTGCTTCAA	CGGGCCGCG	GCCAGATTTC	GTGCCCATGT	CCAGCTCACT	GTGGCAGTTT	1920
TTATCACTT	CAAGTGTGAA	CCAGAGCGTA	CGACTGTGTA	CCAGGGCCAC	ACAGCCCTAC	1980
TGCAATGCGA	GGCCAGGGG	GACCCCAAGC	CGCTGATTCA	GTGAAAGGCG	AAGGACCGCA	2040
TCCTGGAGCC	CACCAAGCTG	GGACCCAGGA	TGCACATCTT	CCAGAAATGG	TCCCTGGTGA	2100



5  
10  
15  
20  
25  
30  
35

```

TCCATGACGT GGCCCTGAG GACTCAGGCC GCTACACCTG CATTGCAGGC AACAGCTGCA 2160
ACATCAAGCA CAGGAGGACC CCCTCTATG TCGTGGACAA GCCTGTGCGG GAGGAGTCGG 2220
AGGGCCCTGG CAGCCCTCCC CCTACAAGA TGATCCAGAC CATTGGGTTG TCGGTGGGTG 2280
CGCTGTGGC CTATCATATT GCGGTGCTGG GCCTCATGTT CTACTGCAAG AAGCGCTGCA 2340
AAGCCAGCGG GCTGCAGAA GAGCCCGAGG GCGAGGAGCC AGAGATGGAA TGCCCTCAACG 2400
GAGGGCTTTT GCAGAACGGG CAGCCCTCAG CAGAGATCCA AGAAGAAGTG GCCTTGACCA 2460
GCTTGGGCTC CGGCCCCGGG GCCACCAACA AACGCCACAG CACAAGTGAT AAGATGCACT 2520
TCCACGCGTC TAGCCTGCAG CCTATCACCA CGCTGGGAA GAGTGAGTTT GGGGAGGTGT 2580
TCCTGGCAAA GGCTCAGGGC TTGGAGGAGG GAGTGGCAGA GACCTTGGTA CTTGTGAAGA 2640
GCCTGCAGAC GAAGGATGAG CAGCAGCAGC TGGACTTCCG GAGGGAGTTG GAGATGTTTG 2700
GGAAGCTGAA CCAAGCCAAC GTGGTGGGCT TCCTGGGCT GTGCCGGAG GCTGAGCCCC 2760
ACTACATGAT GCTGGAATAT GTGATCTGG GAGACCTCAA GCAGTTCTCT AGGATTTCCT 2820
AGAGCAAGGA TGAATAATTG AAGTACAGC CCTCAGCAC CAAGCAGAAG GTGGCCCTAT 2880
GCACCCAGGT AGCCCTGGGC ATGGAGCACC TGTCCAACAA CGCTTTGTG CATAGGACT 2940
TGCTGGCGG TAAGTGCCTG GTCAGTGGCC AGAGACAAGT GAAGTGTCT GCCCTGGGCC 3000
TCAGCAAGGA TTGTACAAAC AGTGAGTACT ACCACTTCCG CCAGGCGCTG GTGCCGCTGC 3060
CTGTGATGTC CCGCAGGCGC ATCTCGGAGG GTGACTTCTC TACCAAGTCT GATGTCTGGG 3120
CCTTCGGTGT GCTGATGTGG GAAGTGTGTA CACATGGAGA GATGCCCAT GGTGGGCGAG 3180
CAGATGATGA AGTACTGGCA GATTTCAGG CTGGGAAGGC TAGACTTCT CAGCCCGAGG 3240
GCTGCCCTTC CAACCTCTAT CGGTGATGC AGCGCTGCTG GGCCCTCAGC CCCAAGGACC 3300
GGCCCTCTCT CAGTGAATTT GCCAGCGCCC TGGGAGACAG CACCGTGGAC AGCAAGCCGT 3360
GAGGAGGAGG CCGCTCAGG ATGGCTGGG CAGGGGAGGA CATCTCTAGA GGGAGCTCA 3420
CAGCATGATG GGCAGATGTC GTGTCTCTCT GGGCTCTGAG GTGCCCTAGT GCAACAGGCA 3480
TTGTGAGGT CTGAGCAGGG CCGTGGCTTT CCTCTCTTC CTCACCTCA TCCTTTGGGA 3540
GGCTGACTTG GACCCAACT GGGCGACTAG GGCTTTGAGC TGGCGAGTTT CCCCTGCCAC 3600
CTCTCTCTCT ATCAGGAGCA GTGTGGTGC CACAGGTAAC CCCAATTCT GGCCTTCAAC 3660
TTCTCCCTCT GACCGGTCCT AACTCTGCCA CTCATCTGCC AACTTTGCTT GGGGAGGGCT 3720
AGGCTTGGGA TGAGCTGGGT TTGTGGGAG TTCTTAATA TTCTCAAGT CTGGGACAC 3780
AGGTTAATG AGTCTCTTG CCACTGGTCC ACTTGGGGT CTAGACCAGG ATTATAGAGG 3840
ACACAGCAAG TGAGTCTCTC CCACTCTGG CTTGTGCACA CTGACCAGA CCCAGCTCT 3900
CCCCACCTTT CTCTCTTTT CTCTCTTAA GTGCTGGCA GATGAAGGAG TTTTCAGGAG 3960
CTTTTGACAC TATATAAAC CCCTTTTGT TATGACCAC GGGCGGCTTT TATATGTAAT 4020
TGACGGTGG GGTGGGTGG CATGGAGGT AGGGGTGGG CCTGGAGATG AGGAGGGTGG 4080
GCCATCTTA CCCACACTT TTATTGTGT CGTTTGTGT TTGTTTGTG TTTTGTGTTT 4140
TGTTTTTGT TTTACACTG CTGCTCTCA TAAATAAGCC TTTTTTA

```

Seq ID NO: 667 Protein sequence

Protein Accession #: NP\_002812

40  
45  
50  
55  
60

```

1 11 21 31 41 51
MGAARGSPAR PRRLPLLSVL LPLLLGGTQT AIVFIKQPSS QDALQRRAL LRCEVEAPGP 60
VHVWLLDGA PVQDTERFPA QGSSLSFAAV DRLQDSGTFQ CVARDDVTGE EARSANASPN 120
IKNIIEAGPV LKHPASEAII PQPTQVTLRC HIDGHFRPTY QWFRDGTPLS DQGSNHTVSS 180
KERNLTLRPA GPEHSLGYSC CAHSAFQOAC SSONFTLSIA DESFARVULA PQDVVVARYE 240
EAMFHCQPSA QPPPSLQWLF EDETPITNRS RPHLRRATV FANGSLLLTQ VRPRNAGIYR 300
CIGQGQRGPP IILEATLHLA EIEDMLFEP RVFTAGSEER VTCLPPKGLP EPSVWVEHAG 360
VRLPTHGRVY QKGHELVLAN IAESDAGVYT CHAANLAGOR RQDVNITVAT VPSWLKKPQD 420
SQLEEGKPGY LDCLTQATPK PTVVWYRNQM LISEDSRPEV FKNGTLRINS VEVDGTWYR 480
QMSSTPAGSI PPARVQVLE KLKFTPPQP QOCMEFDKEA TVPCSATGRE KPTIKWERAD 540
GSSLPEWVTD NAGTLHFAVR TRDDAGNYTC IASNGPQGI RAHVQLTVAV FITFKVEPER 600
TTVYQGHIAL LQCEAQGDPK PLIQWKGKDR ILDTPLKGR MHIFQNGSLV IHDVAPEDSG 660
RYTCLAGNSC NIKHTEAPLY VVDKPVPEES EGPGSPPPYK MIQTIGLSVG AAVAYIIAIVL 720
GLMFYCKKRC KAKRLQKQPE GEEPEMECLN GGFLQNGQPS AEIQEEVALT SLGSGPAATN 780
KRHSTSDKMH FPRESLQPTT TLGKSEFGEV FLAKAQGLEE GVAETLVLVK SLQTKDEQQQ 840
LDPRRELWMP GKLNHANVVR LLGLCREAEP HYMVLEYVDL GDLKQFLRIS KSKDEKLKSO 900
PLSTKQKVAL CTQVALGMEH LNNRPFVHED LAARNCLVSA QRQVKSALG LSKDVYNSFY 960
YHFRQAWVPL RMSPPEAILE GDFSTKSDVM AFGVLMWEVF THGEMPHGGQ ADDEVLDLQ 1020
AGKARLPQPE GCPSKLYRLM QRQWALSPKO RPSFSEIASA LGDSTVDSKP

```

Seq ID NO: 668 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..1389

65  
70  
75  
80  
85

```

1 11 21 31 41 51
ATGGGCTACC AGAGGCAGGA GCCTGTCTAT CCGCCGAGA GAGATTAGA TGACAGAGAA 60
ACCCCTGTTT CTGAACATGA GTATAAGAG AAAACCTGTC AGTCTGTGTC TCTTTTAAAT 120
GTTGTCAACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT 180
GGGTTTCTCT TGGGAATATT GCTTTTATTC TGGGTTTCAT ATGTTACGGA CTTTCCCTT 240
GTTTATTGTA TAAAGAGGAG GGCCCTCTCT GGAACAGATA CCTACCAGTC TTTGGTCAAT 300
AAAACCTTTC GCTTTCAGG GTATCTGCTC CTCTCTGTTT TTCAGTTTTT GTATCCTTTT 360
ATAGCAATGA TAAGTTACAA TATAATAGCT GGAGATACTT TGAGCAAAGT TTTTCAAGA 420
ATCCCGAGG TTGATCTCTA AAACGTGTTT ATTGCTGCCC ACTTCATTAT TGGACTTTCC 480
ACAGTTACTT TTAATCTGCC TTTATCCTTG TACCGAAATA TAGCAAAGCT TGAAGAGGTC 540
TCCCTCATCT CTACAGGTTT AACAACTCTG ATTCTTGGAA TTGTAATGGC AAGGGCAATT 600
TCACTGGGTC CACACATACC AAAAACAGAA GACGCTTGGG TATTGCAAA GCCCAATGCC 660
ATTCAAGCGG TCGGGTTAT GTCTTTTGCA TTTATTGCCC ACCATAACTC CTCTCTAGTT 720
TACAGTTCTC TAGAAGAACC CACAGTAGCT AAGTGGTCCC GCCTTATCCA TATGTCATC 780
GTGATTCTTG TATTATCTTG TATATTCTTT GCTACATGTC GATACTTGAC ATTACTGGC 840
TTCACCAAG GGGACTTATT TGAAATTTAC TGCAGAAATG ATGACCTGGT AACATTGGA 900
AGATTGTTG ATGGTGTGAC TGTCAITTTG ACATACCTTA TGGAAATGCT TGTGACAGA 960
GAGGTAATTG CCAATGTGTT TTTTGGTGGG AATCTTTCAT CGGTTTCCA CATTTGTTGA 1020
ACAGTGTATG TCATCACTGT GCTCTGTGCA ACTCCCTTCA TTTTATCAT TCCATCAGCC 1140
GTTCTAGAAC TCAATGTGTG GCTCTGTGCA ACTCCCTTCA TTTTATCAT TCCATCAGCC 1200
TGTATCTGTA AACTGTCTGA AGAACCAAGG ACACACTCCG ATAAGATTAT GTCTTGTGTC 1260
ATGCTTCCCA TTGGTGTCTG GTGATGTTT TTTGGATTGC TCATGGCTAT TACAAACT

```



CAAGACTGCA CCCATGGGCA GGAAATGTTT TACTGCTTTC CTGACAATTT CTCTCTCACA 1320  
AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACCTT CTACTTTAAA TATTAGTATC 1380  
TTTCAATGA

5 Seq ID NO: 669 Protein sequence  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
10	MGYQRQEPVI	PPQRDLDDRE	TLVSEHEYKE	KTCQSAALFN	VVNSIIGSGI	IGLPYSMKQA	600
	GFPLIGILLF	WVSYVTDPSL	VLILKGGAAS	GTDTYQSLVN	KTFGPPGYLL	LSVLQFLYPP	1200
	IMISYNNIA	GDTLSKVPR	IPGVDPENVF	IGRHFIIGLS	TVTFITPLSL	YRIANLQKGV	1800
	SLISTGLTTL	ILIGVMARAI	SLGPHIKPTE	DAWFAKPNA	IQAVGVMSFA	FICHENSLVF	2400
	YSSLEPTVA	KWSRLHMSI	VISVPICFH	ATCGYLTFTG	FTQGDLFENY	CRNDLLVTFG	3000
15	RCFGYCTVIL	TYMPCFVTR	EVIANVPFG	NLSSVPHIVV	PMQVITVATL	VSLLDICLGI	3600
	VDLNGVQMA	TPLIPIIPSA	CYTKLSEVPT	THSDKIMSCV	MLPIGAVVMV	FGFVMAITNT	4200
	OLECHGOECF	YCFPDNFSLT	NTLSLSSEOT	TQTLTSLISL	FQ		

20 Seq ID NO: 670 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1284

	1	11	21	31	41	51	
25	ATGGGCTAC	AGAGGCAGGA	GCCTGTCATC	CGCCGCAGA	GAGGATTGCC	TTATTCAATG	600
	AAGCAAGCTG	GTTTTCCTTT	GGGAATATTG	CTTTTATTCT	GGGTTCATCA	TGTTACAGAC	120
	TTTTCCTGT	TTTATTATGAT	AAAGAGGGTG	CGCCTCTCTG	GAACACAGATC	TCCAGGCTTG	180
	TGCTCAATA	AAACTTTCCG	CTTTCACAGG	TATCTGCTCC	TCTCTGTTCT	TCAGTTTGTG	240
	TATCCTTTTA	TAGCAATGAT	AGATTACAA	ATAATAGCTG	GAGATACTTT	GACCAAAAGT	300
30	TTTCAAGAAA	TCCACGAGAT	TAACTCTGAA	AACGCTGTTA	TTGTCGCCCA	CTTCATTATT	360
	GGACTTTCCA	CACTTACCTT	TACTCTGCCT	TATCTCTTGT	ACCGAAATAT	ACCAAGCACTT	420
	GGAAAGGTC	CCCTCATCTC	TACAGGTTTA	ACAACTCTGA	TTCTTGGAAT	TGTAATGGCA	480
	AGGGCACTT	CACCTGGGTC	ACACATACCA	AAACACGAGG	ACGCTTGGGT	ATTTCGAAAG	540
	CCCAATGCCA	TTCAACGGCT	CGGGGTATG	TCCTTTGCTT	TTATTGGCCA	CCATACTCCG	600
35	TTCTTAGTTT	ACAGTCTCT	AGAAGAACCC	ACAGTAGCTA	AGTGTGCCG	CCTTATCCAT	660
	ATGTCATCG	TGATTTCGT	ATTATCTGT	ATATTCTTTG	CTACATGTGG	ATACTTGCAC	720
	TTTATCTGCT	TCACCAAGG	GGACTTATT	GAAATATTCT	GCAGAAATGA	TGACCTGGTA	780
	ACATTGGA	GATTTGTGA	TGGTGCTACT	GTCATTTTGA	CATACCCATT	GGAATGCTTT	840
	GTGACAGAG	AGGTAATTGC	CAATGTGTT	TTTGGTGGA	ACTTCTCATC	GGTTTCCAC	900
40	ATTGTTGTAA	CAGTGAATGG	CATCATTGCT	GCCACCGCTG	TGTCATTGCT	GATTGATGCG	960
	CTCGGATAG	TTCTAGAAT	CAATGGTGT	CTCTGGCAA	CTCCCTCAT	TTTTCATTCT	1020
	CCATCAGCCT	GTTATCTGAA	ACTGCTGAA	GAACCAAGGA	CACACTCCGA	TAAAGATTAT	1080
	TTCTGTGTCA	TGCTTCCCAT	TGGTCTGTG	GAGATGTTCT	TAGGATTTCG	TACTGCCTAT	1140
	ACAAATPACT	AAGACTGCAC	CCATGGGCAG	GAAATGTTCT	ACTGCTTTCT	TGACAAATTC	1200
45	TCTCTCACAA	ATACTCACTG	GTCTCATGTT	CAGCAGACAA	CACAACCTTC	TACTTTAAAT	1260
	ATTAGTACT	TTCAACTCGA	GTAA				

Seq ID NO: 671 Protein sequence  
Protein Accession #: Eos sequence

	1	11	21	31	41	51	
	MGYQRQEFVI	PPQRGLPYSM	KQAGFPLGIL	LLFWVSVD	PSLVLLIKGG	ALSGTDITYQS	60
55	LVNKTGPFPG	YLLLSVLQFL	YFPIAMISYI	IJAGDTLSRV	QRIQPGVDE	NVFIQRHPTI	120
	GLSTVTTFYI	LSLYRNIACL	KGVSILSTGL	TLILGVMA	RAILSGHPH	KTEDAWNFAK	180
	PNAIQAVGVM	SFAFICHHNS	FLWVSSLEEP	TVAKWSRLH	MSIVISVPIC	IFFGATCGYLT	240
	FTVTGQDGLF	ENYCRNDDL	TGPRFCYGVY	VILTYMECF	TVREIANVF	PGCNLSSVFR	300
	IVTMTVMYI	ATIVSLLLDC	LSGIVLELNG	LCATPLFPI	SPCYKLKSE	EPRTSHSKIM	360
60	SCVMPLIGAV	VMVPGFVMAI	TNTQDCTHQG	EMFYCPEPNF	SLNTNTESEV	QQTQLTSLT	420
	ISIFOLE						

Seq ID NO: 672 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1203

	1	11	21	31	41	51	
	ATGGGCTACC	AGAGGCGAGGA	GCGTGTATC	CGCGCGCAGT	TTTCCCTTGT	TTTATTGATA	60
	AAAGAGGGGG	COCCTCTCGG	AACAGATATC	TACCAGCTCT	TGGTCAATTA	AACCTTTCGGC	120
70	TTTCCAGGCT	ATCTGCTCCT	CTCTGTGTTT	CAGTTTGTGT	ATCTTTTAT	AGCAATGATA	180
	AGTTACAATA	TAATAGCTGG	AGATACTTTG	AGCAAAGTTT	TTCAAAGAAT	CCGAGGAGTT	240
	GATCCTGAAA	ACGTGTGTTT	TGGTGGCCAC	TTCAATTATG	GACTTTCAC	GTTTACCTTT	300
	ACTCTGCCTT	TATCTCTGTA	CGCAAAATATA	GCAAAAGCTG	GAAAGGCTCT	CCTTACTCTC	360
75	ACAGGTTTAA	CAACTCTGAT	CTCTGGAAT	GTAATGGCAA	GGGCAATTTC	ATCGGGGTCA	420
	CACATACCAA	CAGCAAGAAG	CGCTTGGGTA	TTGCGAAGC	CCATGCCATT	TCAGAGCCGT	480
	GGGGTTATGT	CTTTTGCATT	TATTTGCCAC	CATAACTCCT	TCTTAGTTTA	CAGTTTCTCA	540
	GAAAGAACCA	CAGTAGCTAA	GTGGTCCCGC	CTTATCCATA	TGTCCATCGT	GATTTCTGTA	600
	TTTATCTGTA	TATCTTTGCG	TACATGTGGA	TACTTGACAT	TTACTGGCTT	CACCCAAGGG	660
80	CAGTATTATT	AAAAATTACT	CAGAAATGAT	CAGCTGTGTA	GATTTGGAA	ATTTTGTGTA	720
	GGTGTCACTG	TAAATTTGAC	ATACCTTATG	GAA'TGCTTG	TCGACAAGAA	GGTATATGCC	780
	AATGTGTTTT	TGGTGGGAAA	CTTTCATCGT	GTTTTCACCA	TTGTGTAAAC	AGTATGTCCT	840
	ATCAGTGTGT	CCACGCTTGT	GTCAATTGCT	ATGATTGCC	TCGGGATATG	TCAGAACTC	900
	AATGGTGTGC	TCGTGCAAC	TCCCTTCATT	TTTATCATTC	CATCGGCTGT	GTTTCTGAAA	960
	CTGTCTGAAG	AACCRAGGAC	ACACTTCGAT	AAGATTATGT	CTTGCTGCAT	GTCTCCCAT	1020
85	GGTGTCTGGT	TGATGGTTTT	TGGATTTCGT	ATGGCTATTA	CAAACTACTA	AGACTGCAAC	1080
	CATGGGCGAC	AAATGTTCTA	CTGCTTTTCT	GACAAATTTCT	CTCTCACAAA	TACCTTCAGT	1140
	TCTCATGTTT	CAGACACAA	CAACCTTTCT	ACTTTAAATA	TTAGTATCTT	TCAACTCGAG	1200

TAA

Seq ID NO: 673 Protein sequence  
Protein Accession #: Eos sequence

5  
10  
15

1	11	21	31	41	51	
MGYQRQEPVI	PPQPSLVLLI	KGGALSGTDT	YQSLVNKTFG	FPGYLLLSVL	QFLYPFIAMI	60
SYNIIAGDTL	SKVFQRIQGV	DPENVFIGRE	FIIGLSTVTF	TLPLSLYRNI	AKLGKVSLLS	120
TGLTLLILGI	VMARAIISLGP	HIPKTEDAMV	FAKPNAIQAV	GVMSFAPICH	HNSFLVYSSL	180
EEPTVAKWSR	LIHMSIVISV	PICIFFATCG	YLTPGTGTG	DLFENYCRND	DLVTGFRFCY	240
GVTVILTYPM	ECFVTRVIA	NVFFGGNLS	VFHIVVTVMV	ITVATLVSL	IDCLGIVLEL	300
NGVLCATPLI	FIIPSACYLK	LSEEPRTSD	KIMSCVMLPI	GAUVMVFGPV	MAINTNQDCT	360
HQEMFYCFP	DNFSLTNTSE	SHVQQTQLS	TLNISIFQLE			

Seq ID NO: 674 DNA sequence  
Nucleic Acid Accession #: Eos sequence  
Coding sequence: 1..1140

20  
25  
30  
35  
40

1	11	21	31	41	51	
ATGGGCTACC	AGAGGCAGGA	GCGTGTATC	CGCCGCGAGG	TCAATAAAC	TTTCGGCTTT	60
CCAGGGATAC	TGCTCCTCTC	TGTTCTTCAG	TTTTGTATC	CTTTTATAGC	AATGATAAGT	120
TACAATATAA	TAGCTGGAGA	TACITTGAGC	AAAGTTTTTC	AAAGAATCCC	AGGAGTTGAT	180
CCTGAAACAG	TGTTTATGG	TCGCCACTTC	ATTATTGGAC	TTTCCACAGT	TACCTTTACT	240
CTGCCCTTAT	CTTTGTACCG	AAATATAGCA	AAGCTTGGAA	AGGTCTCCCT	CATCTCTACA	300
GGTTTAACAA	CTCTGATTCT	TGGAATTGTA	ATGGCAAGGG	CAATTTCAC	GGGTCCACAC	360
ATACCAAAAA	CAGAAGACGC	TTGGGTATTT	GCAAGGCCCA	ATGCCATTCA	AGCGGTGGG	420
GTATGTCTTT	TGCAATTTAT	TTGCCACCAT	AACTCCTTCT	TAGTTTACAG	TTCTCTAGAA	480
GAACCCACAG	TAGCTAAGTG	GTCCCGCCTT	ATCCATATGT	CCATCGTGAT	TTCTGTATTT	540
ATCTGTATAT	TCTTTGCTAC	ATGTGGATAC	TTGACATTTA	CTGGCTTCAC	CCAAGGGGAC	600
TTATTGAAAA	ATTACTGCAG	AAATGATGAC	CTGGTAACAT	TTGGAAGATT	TTGTATAGGT	660
GTCACTGTCA	TTTTGACATA	CCCTATGGAA	TGCTTTGTGA	CAAGAGAGGT	AATTGCCAAT	720
GTGTTTTTGG	GTGGGAATCT	TTCACTGGTT	TTCCACATTG	TTGTAACAGT	GATGGTCATC	780
ACTGTAGCCA	CGCTTGTGTC	ATTGCTGATT	GATTGCCTCG	GGATAGTTCT	AGAACTCAAT	840
GGTGTGCTCT	GTGCAACTCC	CCTCATTTTT	ATCATTCAT	CAGCCTGTTA	TCTGAAACTG	900
TCGGAAGAAC	CAAGGACACA	CTCCGATAAG	ATTATGTCTT	GTGTGATGCT	TCCCATTTGT	960
CTGTGTGGTA	TGTTTGTGG	ATTGCTCATG	GCTATTACAA	ATACTAAGA	CTGCAACCAT	1020
GGGCAGGAAA	TGTTCTACTG	CTTTCCTGAC	AATTCTCTTC	TCACAAATAC	CTCAGAGTCT	1080
CATGTTTCAGC	AGACACACAC	ACTTCTACT	TTAAATATTA	GTATCTTTCA	ACTCGAGTAA	

Seq ID NO: 675 Protein sequence  
Protein Accession #: Eos sequence

45  
50  
55

1	11	21	31	41	51	
MGYQRQEPVI	PPQVKNKTFG	PGYLLLSVLQ	FLYPFIAMIS	YNIAGDTLS	KVFQRIQGV	60
PRNVFIGREF	IIGLSTVTF	LPLSLYRNIA	KLKGVSLIST	GLTLLILGIV	MARAIISLGP	120
IPKTEDAMV	AKPNAIQAV	VMSFAPICH	NSFLVYSSLE	EPTVAKWSR	IHMSIVISV	180
ICIFFATCGY	LTPGTGTG	LFENYCRND	LVTGFRFCY	VTIVTYPM	CFVTRVIAN	240
VFFGGNLSV	FHIVVTVMV	TVATLVSL	DCLGIVLEL	GVLCATPLIF	IIPSACYLK	300
SEEPRTSDK	IMSCVMLPI	AVVMVFGPV	AITNTQDCTH	GQEMFYCFPD	NFSLTNTSES	360
HVQQTQLST	LNISIFQLE					

Seq ID NO: 676 DNA sequence  
Nucleic Acid Accession #: NM\_006853.1  
Coding sequence: 26..874

60  
65  
70  
75  
80

1	11	21	31	41	51	
AGGAATCTGC	GCTCGGTTTC	CGCAGATGCA	GAGGTTGAGG	TGGCTGCGGG	ACTGGAAGTC	60
ATCGGGCAGA	GGTCTCACAG	CAGCCAAGGA	ACCTGGGGCC	CGCTCCTCCC	CCCTCCAGGC	120
CATGAGGATT	CTGCAGTTAA	TCCTGCTTGC	TCTGGCAACA	GGGCTGTAG	GGGGAGAGAC	180
CAGGATCATC	AAGGGGTTCC	AGTGCAAGCC	TCACTCCACG	CCCTGGCAGG	CAGCCCTGTT	240
CGAGAAGAGC	CGGCTACTCT	GTGGGGGAC	GCTCATCGCC	CCCAGATGGC	TCTGACAGC	300
AGCCCACTGC	CTCAAGCCCC	GCTACATAGT	TCACCTGGGG	CAGCACAACC	TCCAGAAGGA	360
GGAGGGCTGT	GAGCAGACCC	GGACAGCCAC	TGAGTCTTTC	CCCCACCCCG	GCTTCAACAA	420
CAGCTCCCC	AACAAAGACC	ACCGCAATGA	CATCATGCTG	GTGAAGATGG	CATCGCCAGT	480
CTCCATCACC	TGGGCTGTGC	GACCCCTCAC	CCTCTCCTCA	CGCTGTGTCA	CTGCTGGCAC	540
CAGCTGCCTC	ATTTCGGGCT	GGGGCAGCAC	GTCCAGCCCC	CAGTTACGCC	TGCTTCACAC	600
CTTGCGATGC	GCCAACATCA	CCATCATTTA	GCACCAGAG	TGTGAGAACG	CCTACCCCGG	660
CAACATCACA	GACACCATGG	TGTGTGCCAG	CGTGACAGAA	GGGGGCAAGG	ACTCCTGCCA	720
GGGTGACTCC	GGGGGCCCTC	TGGTCTGTAA	CCAGTCTCTT	CAAGGCATTA	TCTCTGGGG	780
CCAGGATCCG	TGTGCGATCA	CCCGAAAGCC	TGGTGTCTAC	ACGAAAGTCT	GCAAAATATG	840
GGACTGGATC	CAGGAGACGA	TGAAGAACAA	TTAGACTGGA	CCCACCCACC	ACAGCCCATC	900
ACCTCCATAT	TCCACTTGGT	GTITGGTTCC	TGTTCACTCT	GTTAATAAGA	AACCCCTAAG	960
CAAGACCTTC	TACGAACATT	CTTTGGGCCT	CCTGGACTAC	AGGAGATGCT	GTCACTTAAT	1020
AATCAACCTG	GGGTTGAAA	TCAGTGAGAC	CTGGATTCAA	ATTCTGCCTT	GAAATATTGT	1080
GACTCTGGGA	ATGACAACAC	CTGTTTGT	CTCTGTTGTA	TCCCCAGCCC	CAAAGACAGC	1140
TCCTGGCCAT	ATATCAAGGT	TTCAATAAAT	ATTGCTAAA	TGAGTG		

Seq ID NO: 677 Protein sequence  
Protein Accession #: NP\_006844.1

85

1	11	21	31	41	51	
MRILQLILLA	LATGLVGET	RIIKGFECFP	HSQFWQAALF	EKTRLLCGAT	LIAPRWLLTA	60

AHCLKPRIV HLQHQHLOKE EGCEQRTAT ESFPHPGFNN SLFNKDHEND IMLVKMASPV 120  
 SITWAVRPLT LSSRCVYAGT SCLISGWGST SSPQLRLPHT LRCANITIIIE HOKCENAYPG 180  
 NITDTMVCAS VQEGGKDSQ GDSSGGLVCN QSLQGIISWG QDPCAITRKP GVYTKVKYV 240  
 DWIQETMKN

Seq ID NO: 678 DNA sequence  
 Nucleic Acid Accession #: Eos sequence  
 Coding sequence: 1..933

1 11 21 31 41 51  
 ATGTGCAGCA ATGGAAGGTG CATCCCGGGC GCCTGGCAGT GTGACGGGCT GCCTGACTGC 60  
 TTCGACAAGA GTGATGAGAA GGAGTGCCCC AAGGCTAAGT CGAAATGTGG CCCGACCTTC 120  
 TTCCCTGTGT CCAGCGGCAT CCATTGCATC ATTGGTCCGT TCCGCTGCAA TGGGTTTGAG 180  
 GACTGTCCCG ATGGCAGCGA TGAAGAGAAC TGCACAGCAA ACCCTCTGCT TTGCTCCACC 240  
 GCGCGCTACC ACTGCAAGAA CGGCCTCTGT ATTGACAAGA GCTTCTCTGT CGATGGACAG 300  
 AATAACTGTC AAGACAACAG TGATGAGGAA AGCTGTGAAA GTTCTCAAGA ACCCGGCACT 360  
 GGGCAGGTGT TTGTGACTTC AGAGAACCAG CTGTGTATT ACCCCAGCAT CACCTATGCC 420  
 ATCATCGGCA GCTCGTCAT TTTTGTGCTG GTGGTGGCCC TGCTGGCACT GGTCTTGAC 480  
 CACCAGCGGA AGCGGAACAA CCTCATGACG CTGCCCGTGC ACCGGCTGCA GCACCCCTGTG 540  
 CTGCTGTCCC GCCTGGTGGT CCTGAGCCAC CCCCAACCAT GCAACGTCAC CTACCAACGTC 600  
 AATAATGGCA TCCAGTATGT GGCCAGCCAG CGCGAGCAGA ATGCGTGGGA AGTAGGCTCC 660  
 CCACCCCTCT ACTCCGAGGC CTTGTGGAC CAGAGGCGCT CGTGGTATGA CCTTCCTCCA 720  
 CGGCCCTACT CTTCTGACAC GGAATCTCTG AACCAGCCCG ACCTGCCGCC CTACCCCTCC 780  
 CGGTCCGGGA GTGCCAACAG TGCCAGCTCC CAGGCAGCCA CGACCTCTCT GAGCGTGGAA 840  
 GACACCAGCC ACAGCCCGGG GCAGCCTGGC CCCCAAGAGG GCACTGCTGA GCCCAGGGAC 900  
 TCTGAGCCCA GCCAGGGCAC TGAAGAAGTA TAA

Seq ID NO: 679 Protein sequence  
 Protein Accession #: Eos sequence

1 11 21 31 41 51  
 MCSNGRCIPG AWQCDGLPDC FDKSDEKECP KAKSKCGPTF FPCASGIHCI IGRFRONGFE 60  
 DCPDGSDEEN CTANPLLCST ARYHCKRNLG IDKSFICDQ NNCQDNDSEE SCRSSQEPGS 120  
 GQVFTVSENQ LVVYPSITYA IIGSSVIFVL VVALLALVLH HQRKRNLMT LPVRLQHPV 180  
 LLSRLVLDH PEHCNVTYNV NNGIQYVASS ABQNASEVGS PPSYSEALLD QRPAYWDLPP 240  
 PPYSDTESL NQADLPYRS RSGSANGASS QAASSLLSVE DTSESPGQPG PQEGTAEPRD 300  
 SEPSQGTEEV

Seq ID NO: 680 DNA sequence  
 Nucleic Acid Accession #: S78203.1  
 Coding sequence: 1..2190

1 11 21 31 41 51  
 ATGAATCCTT TCAGAAAAA TGAGTCCAAG GAAACTCTTT TTTCACCTGT CTCCATTGAA 60  
 GAGGTACCAC CTCGACCACC TAGCCCTCCA AAGAAGCCAT CTCGACAAT CTGTGGCTCC 120  
 AACTATCCAC TGAGCATGTG CTTCAITGTG GTGAATGAAT TCTGCGAGCG CTTTCTCTAT 180  
 TATGGAATGA AAGCTGTGCT GATCCTGTAT TTCTGTATT TCTGCACTG GAATGAAGAT 240  
 ACCCTCCATC CTATATACCA TGCCCTCAGC AGCCTCTGTT ATTTTACTCC CATCCTGGGA 300  
 GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTCAAGACAA TCATCTATCT CTCTTGCTG 360  
 TATGTGCTTG GCCATGTGAT CAAGTCCTTG GGTGCCTTAC CAATACTGGG AGGACAAGTG 420  
 GTACACACAG TCCATCATTT GATCGGCTTG AGTCTAATAG CTTTGGGGAG AGGAGGCATC 480  
 AAACCCCTGT TGGCAGCTTT TGGTGGAGAC CAGTTTGAAG AAAACATGC AGAGGAACGG 540  
 ACTAGATACT TCTCAGCTTT CTACCTGTCC ATCAATGCAG GGAGCTTGAT TTCTACATT 600  
 ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGGAG AAGACTGCTA TGCATTGGCT 660  
 TTTGGAGTTC CAGGACTGCT CATGTGAATT GCACTTGTG TGTTTGCAAT GGGAAAGCAA 720  
 ATATACAAATA AACCAACCCC TGAAGGAAAC ATAGTGGCTC AAGTTTTCAA ATGTATCTGG 780  
 TTTGCTATTT CCAATCGTTT CAAGAACCCT TCTGGAGACA TTCCAAAGCG ACAGCACTGG 840  
 CTAGACTGGG CAGCTGAGAA ATATCCAAAG CAGCTCAITTA TGGATGTAAG GGCAGTGACC 900  
 AGGGTACTAT TCCCTTATAT CCCATTGCCC ATGTTCTGGG CTCTTTTGA TCAGCAGGGT 960  
 TCACGATGGA CTTTGCAAGC CATCAGGATG AATAGGAATT TGGGGTTTT TGTGCTTAC 1020  
 CCGGACCAGA TGCAGGTTCT AAATCCCTTT CTGCTTCTTA TCTTCATCCC GTTGTGTGAC 1080  
 TTTGTCAATT ATGCTCTGGT CTCCAAGTGT GGAATTAACT TCTCATCACT TAGGAAAATG 1140  
 GCTGTTGGTA TGATCCTAGC GTGCCCTGGCA TTGCAAGTTG CGGCAGCTGT AGAGATAAAA 1200  
 ATAAATGAAA TGGCCCCAGC CCAGTCAGGT CCCAGSAGG TTTTCTACA AGTCTTGAAT 1260  
 CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTC TCTGTTGATA 1320  
 GAGTCCATCA AATCCTTTCA GAAAACACCA CACTATTCCA AACTGCACCT GAAAACAAA 1380  
 AGCCAGGATT TTCACTTCCA CCGTAAATAT CACAATTTGT CTCTCTACAC TGAGCATTTCT 1440  
 GTGCAGGAGA AGAATGTGTA CAGTCTGTG ATTGTGAAG ATGGGAACAG TATCTCCAGC 1500  
 ATGATGTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGGT 1560  
 AACACTTTGC ATAAAGATGT CAACATCTCC CTGAGTACAG ATACCTCTCT CAATGTTGGT 1620  
 GAAGACTATG GTGTCTTGC TTATAGAAT GTGCAAGAG GAGAATACCC TGCAGTGCAC 1680  
 TGTAGAACAG AAGATAAGAA CTTTCTCTG AATTGGGTC TTCTAGACTT TGTGTGACGA 1740  
 TATCTGTTTG TTATTACTAA TAACACCAAT CAGGGTCTTC AGGCCTGGAA GATTGAAGAC 1800  
 ATTCCAGCCA ACAAAATGTC CATTGGCTGG CAGCTACCAC AATATGCCCT GGTACAGCT 1860  
 GGGGAGGTCA TGTCTCTGT CACAGTCTTT GAGTTTTCTT ATTCTCAGC TCCCTTAGC 1920  
 ATGAAATCTG TGCTCCAGCG AGCTTGGCTA TTGACAATTG CAGTTGGGAA TATCATCTG 1980  
 CTGTTGTGGT CACAGTTCAG TGGCCTGGTA CAGTGGGCG AATTCAATTT GTTTCTCTG 2040  
 CTCCTGCTGG TGATCTGCTC GATCTTCTCC ATCATGGGCT ACTACTATGT TCCTGTAAAG 2100  
 ACAGAGGATA TGGCGGGTCC AGCAGATAAG CACATTCCTC ACATCCAGGG GAACATGATC 2160  
 AAACATAGAG CCAAGAGAC AAAACTCTGA

Seq ID NO: 681 Protein sequence  
 Protein Accession #: AAB34388.1

1 11 21 31 41 51  
MNPFOKNESEK ETLFSPVSIE EVPPRPSPSP KKPSTICGS NYPLSIAFIV VNEFCERFSY 60  
YGMKAVLILY PLYPLHWNED TSTSIYHAFS SLCYFTPILG AALADSWLKG FKTIYLSLV 120  
5 YVLGHVIKSL GALPILGGQV VHTVLSLIGL SLIALGTGGI KPCVAAFQGD QPEEKHABER 180  
TRYFSVYPLS INAGSLISTF ITPMLRGDVQ CFGEDCYALA FGVPGLLMVI ALVVPMGSK 240  
IYNKPPPEGN IVAQVFKCIW FAISNRFPNR SGDIPKQHW LDWAAEKYPK QLIMDVKALT 300  
RVFLYIPLF MFALLDQCG SRWTLQAIM NRNLGPPVLQ PDQMQLNPF LVLIPIPLFD 360  
FVIYRLVSKC GINFFSLRKM AVGMILACLA FAVAAVEIK INEMAPAQSG PQEVFLQVLN 420  
10 LADDEVKVTV VGNENNELLI ESIKSFQKTP HYSKLHLKTK SDFHFLPKY HNLSTLYEHS 480  
VQERNWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG 540  
EDYGVSAIRT VQGEYPAVH CRTEDKNPSL NLGLLDFGAA YLFVITNNNT QGLQAWKIED 600  
IFANKMSIAW QLPQYALVTA GEVMSFVTGL EFSYSQAPSS MKSVLQAANL LTIAGVNIIV 660  
15 LVVAQFSGLV QWAEFILFSC LLLVICLIFS IMGYVYPVK TEDMRGPADK HIPHIQGNMI 720  
KLETKTKKL

Seq ID NO: 682 DNA sequence  
Nucleic Acid Accession #: NM\_016077.1  
Coding sequence: 128..667

1 11 21 31 41 51  
TCGCTTTGTG ATTCTTGATC CGGAACCTTG TCACCCAGGA ACCCCGGAAG AGGTAGCTCA 60  
CGCGATAGAA ACGTGTTCGC TTGCCAGAA GAAGGGAAGG CGCGAGTGAG GAAAGGAGGT 120  
25 ACTGTAGATG CCTCCAAAT CCTTGGTTAT GGAATATTG GCTCATCCA GTACACTCGG 180  
CTTGGCTGTT GAGATTGCTT GTGGCATGTG CTGGGCTGG AGCCTTCGAG TATGCTTTGG 240  
GATGCTCCCC AAAAGCAAGC CGAGCAAGAC ACACACAGAT ACTGAAAGTG AAGCAAGCAT 300  
CTTGGGAGAC AGCGGGGAGT ACAAGATGAT TCTTGTGGTT CGAAATGACT TAAAGATGGG 360  
30 AAAAGGGAAG GTGGCTGCCC AGTGTCTCA TGCTGTGTT TCAGCCTACA AGCAGATTCA 420  
AAGAAGAAAT CCTGAAATGC TCAAACAATG GGAATACTGT GCGCAGCCCA AGGTGGTGGT 480  
CAAGCTCCT GATGAAGAAA CCTGTATTGC ATTATTGGCT CATGCAAAAA TGCTGGGACT 540  
GACTGTAAGT TTAATTCAAG ATGCTGGACG TACTCAGATT GCACCAGGCT CTCAAACTGT 600  
CCTAGGGATT GGGCCAGGAC CAGCAGACCT AATTGACAAA GTCACTGGTC ACCTAAACT 660  
35 TTACTAGTGT GACTTTGATA TGACAACAAC CCTCCATCA CAAGTGTGTT AAGCCTGTCA 720  
GATTCTAACA ACAAAAGCTG AATTCTTCA CCCAACTTAA ATGTTCTTGA GATGAAAATA 780  
AAACCTATTG CCATGTTCTA AAAAA

Seq ID NO: 683 Protein sequence  
Protein Accession #: NP\_057161.1

1 11 21 31 41 51  
MPSKSLVMEY LAHPSTLGLA VGVACGMCLG WSLRVCFGML PKSKTSKTHT DTSEASILG 60  
45 DSGEYKMLLV VRNDLKMKGK KVAACQSHAA VSAYKQIQR NPMLKQWEY CQPKVVKKA 120  
PDEETLIALL AHAKMLGLTV SLIQDAGRTO IAPGSQTVLG IGPSPADLID KVTGHLKLY

Seq ID NO: 684 DNA sequence  
Nucleic Acid Accession #: NM\_004864.1  
Coding sequence: 26..952

1 11 21 31 41 51  
CGGAACGAGG GCAACCTGCA CAGCCATGCC CGGGCAAGAA CTCAGGACGG TGAATGGCTC 60  
55 TCAGATGCTC CTGGTGTGTC TGCTGCTCTC GTGGCTGCGG CATGGGGGCG CCTGTCTCT 120  
GCGCGAGGCG AGCGCGGCAA GTTCCCGGG ACCCTCAGAG TTGCACTCCG AAGACTCCAG 180  
ATTCCGAGAG TTGCGGAAC GCTACGAGGA CCTGCTAACC AGGCTGCGGG CCAACCCAGG 240  
CTGGGAAGAT TCGAACCCG ACCTCGTCCC GCGCCCTGCA GTCCGGATAC TCAGCCAGGA 300  
AGTGCGGCTG GATCCGCGG GCGACCTGCA CTGCGTATC TCTCGGGCGG CCTTCCCGA 360  
60 GGGGCTCCCC GAGGCTCCCC GCCTTCACCG GGCTCTGTTC CGGCTGTCCC CGACGGCGTC 420  
AAGGTGCTGG GAGCTGACAC GACCGCTGCG GCCTCAGCTC AGCCTTGCAA GACCCCAAGC 480  
CGCCGCGCTG CACTCCGAC TGTGCGCGCC GCGTGCAG TCGGACCAAC TGCTGCGAGA 540  
ATCTTGCTCC GCACGGCCCC AGCTGGAGTT GCACTTGCGG CCGCAAGCCG CCAGGGGGCG 600  
CCGCAGAGCG CGTGCGCGCA ACGGGGACGA CTGTCCGCTC GGGCCCGGGC GTTGTGCGG 660  
65 TCTGCACAGG GTCCGCGCTG CGCTGGAAGA CTGGGCTGG GCGGATTGGG TGCTGTGCGC 720  
ACGGGAGGTG CAAGTGACCA TGTGCTATCG CGCTGCGCG AGCCAGTTCG GGGCGGCAAA 780  
CATGCACGCG CAGATCAAGA CGAGCCTGCA CGCCTGAAG CCCGACACGG AGCCAGCGCC 840  
CTGCTGCGTG CCGCCAGCT ACAATCCCAT GGTGCTCAT CAAAAGACCG ACACCGGGGT 900  
GTGCTCCAG ACCTATGATG ACTTGTAGC CAAAGACTGC CACTGCATAT GAGCAGTCCT 960  
70 GGTCTCTCCA CTGTGCACCT GCGCGGGGGA GCGACCTCA GTTGTCTGCG CCTGTGGAAT 1020  
GGGCTCAAGG TTCCTGAGAC ACCGATTCC TGCCCAACA GCTGTATTTA TATAAGTCTG 1080  
TTATTTATTA TTAATTTATT GGGGTGACCT TCTTGGGAGC TCGGGGCGTG GTCTGATGGA 1140  
ACTGTGTATT TATTTAAAC TCTGGTGATA AAAATAAAGC TGCTGAACT GTTAAAAAAA 1200  
AAAA

Seq ID NO: 685 Protein sequence  
Protein Accession #: NP\_004855.1

1 11 21 31 41 51  
80 MPQGELETVN GSQMLLVLLV LSWLPHGGAL SLAEASRAF PGPSELHSED SRPRELRKY 60  
EDLLTRLRAN QSWEDSNTOL VPAPAVRILT PEVRLGSGGH LHLRISRAAL PEGLEPEASRL 120  
HRALFRLSPT ASRSWVTRP LRRQLSLARP QAPALHLRLS PPSQSDQLL AESSARPOL 180  
ELHLRPQAR GRRRARARNG DDCLGPGRC RLHTVRASL EDLGMADWVL SPREVQVTCM 240  
85 IGACPSQFRA ANMHAQIKTS LHLKPDTEP APCCPASYN PMVLQKTDI GVSILQTYDDL 300  
LAKDCHCI

Seq ID NO: 686 DNA sequence

Nucleic Acid Accession #: NM\_002423.2

Coding sequence: 48..851

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      ACCAAATCAA CCATAGGTCC AAGAACAATT GTCTCTGGAC GGCAGCTATG CGACTCACCG 60
      TGGTGTGTGC TGTGTGCTCG CTGCTCTGGCA GCCTGGCCCTT GCGCTGCTCT CAGGAGGCGCG 120
      GAGGCATGAG TGAGCTACAG TGGGAACAGG CTCAGGACTA TCTCAAGAGA TTTTATCTCT 180
10     ATGACTCAGA AACAAAAAAT GCCAACAGTT TAGAAGCCAA ACTCAAGGAG ATGCAAAAAAT 240
      TCTTTGGGCT ACCTATAACT GGAATGTTAA ACTCCGCGT CATAGAAATA ATGCAGAAGC 300
      CCAGATGTGG AGTGCCAGAT GTTGCAGAAT ACTCACTATT TCCAAATAGC CCAAAATGGA 360
      CTTCCAAAGT GGTCACTAC AGGATCGTAT CATATACTCG AGACTTACCG CATATTACAG 420
      TGGATCGATT AGTGTCAAAG GCTTTAAACA TGTGGGGCAA AGAGATCCCT CTGCATTTC 480
      GGAAGATTGT ATGGGGAATC GCTGACATCA TGATTGGCTT TGCCTGAGGA GCTCATGGGG 540
15     ACTCTACCC ATTGTATGGG CCAGGAAACA CGCTGGCTCA TGCCTTTGCG CCTGGGACAG 600
      GTCTCGGAGG AGATGCTCAC TTCGATGAGG ATGAACGCTG GACGGATGGT AGCAGTCTAG 660
      GGAATTAACCT CCTGTATGCT GCAACTCATG AACTTGGCCA TTCTTTGGGT ATGGGACATT 720
      CCTCTGATCC TAATGCAGTG ATGTATCCAA CCTATGAAA TGGAGATCCC CAAAATTTTA 780
      AACTTTCCCA GGAATGATAT AAAGGCATTC AGAACTATA TGGAAAGAGA AGTAATTCAA 840
20     GAAAGAAATA GAACTATCAG GCAGAACATC CATTCAATCA TTCAATGGAT TGTATATCAT 900
      TGTTCACAAA TCAGAAATGA TAAGCACTGT TCCTCCATCT CATTAGCAAA TTATGTCACC 960
      CTTTTATTAT GCAGTTGGTT TTTGAATGTC TTTCACTCCT TTTATTGGTT AAATCCTTT 1020
      ATGGTGTGAC TGTGCTTAT TCCATCTATG AGCTTTGTCA GTGCGCGTAG ATGTCAATAA 1080
25     ATGTTACATA CACAAATAAA TAAATGTTT ATTCCATGTT AATTTTA

```

Seq ID NO: 687 Protein sequence

Protein Accession #: NP\_002414.1

```

30     1      11      21      31      41      51
      |      |      |      |      |      |
      MRLTVLCVAV LLPGLALPL PQEAGGMSEL QWEQAQDYLK RPYLYDSETK NANSLEAKLK 60
      EMQKFFGLPI TGMNLSRVIE IMQKPRCGVP DVAEYSLFPN SPKWTSKVVT YRIVSYTRDL 120
      PHITVDRIVS KALNMWGKEI PLHFRKVVWG TADIMIGPAR GAHGDSYFPD GPGNTLAHAF 180
35     APGTGLGSDA HFDEDERWTD GSSLGINFLY AATHELGHSL GMGHSSDPNA VMYPYTGND 240
      PQNFKLSQDD IKGIQKLYGK RENSRRK

```

Seq ID NO: 688 DNA sequence

Nucleic Acid Accession #: NM\_005221.3

Coding sequence: 1..870

```

40     1      11      21      31      41      51
      |      |      |      |      |      |
      ATGACAGGAG TGTTTGACAG AAGGGTCCCC AGCATCCGAT CCGGCGACTT CCAAGCTCCG 60
      TTCACAGCGT CCGCAGCTAT GCACCATCCG TCTCAGGAAT CCGCACTTT GCCCGAGTCT 120
45     TCAGCTACCG ATTCTGACTA CTACAGCCCT ACGGGGGGAG CCGCGCAGG CTACTGCTCT 180
      CCTACCTCGG CTTCCTATGG CAAAGCTCTC AACCCCTACC AGTATCAGTA TCACGGGCTG 240
      AACGGCTCCG CCGGGAGCTA CCCAGCCAAA GCTTATGCCG ACTATAGCTA CGCTAGCTCC 300
      TACCACCACT ACGGCGGCGC CTACAACCGC GTCCCAAGCG CCACCAACCA GCCAGAGAAA 360
      GAASTGACCG AGCCCGAGGT GAGAAATGGT AATGGCAAA CAAAGAAAGT TCGTAAACCC 420
50     AGGACTATT ATTCCAGCTT TCAGCTGGCC GCATTACAGA GAAGGTTTCA GAAGACTCAG 480
      TACTCGCTCT TGCCGGAACG CGCCGAGCTG GCGGCTCGC TGGGATTGAC ACAACACAG 540
      GTGAAAATCT GGTTCAGAA CAAAAGATCC AAGATCAAGA AGATCATGAA AAACGGGGAG 600
      ATGCCCCCGG AGCACAGTCC CAGCTCCAGC GACCCAATGG CGTGTAACTC GCCGCGTCT 660
      CCAGCGGTGT GGGAGCCCCA GGGCTCGTCC CGCTCGCTCA GCCACCAACC TCATGCCAC 720
55     CCTCCGACCT CCAACCAATC CCCAGCGTCC AGCTACCTGG AGAACTCTGC ATCCTGGTAC 780
      ACAAGTGCAG CCAGTCAAT CAATTCACAC CTGCCGCGCG CGGCTCCTT ACAGCACCCG 840
      CTGGCGCTGG CCTCGGGAC ACTCTATTAG

```

Seq ID NO: 689 Protein sequence

Protein Accession #: NP\_005212.1

```

60     1      11      21      31      41      51
      |      |      |      |      |      |
      MTGVFDRRVP SIRSGDFQAP FQTSAMHHP SQESPTLPES SATDSDYSP TGGAPHGYCS 60
      PFSASYGKAL NPYQVQYHGV NGSAGSYPK AYADYSYASS YHQYGGAYNR VPSATNQPEK 120
      EVTEPEVRNV NGKPKVKRP RTIYSSFLA ALQRRPQKTQ YLALPERAEL AASLGLTQTQ 180
      VKIWFQNRKS KIKKIMKNGE MPPEHSPSSS DPMACNSPQS PAVWEPQSSS RSLSHHPAH 240
      PPTSNSPAS SYLENSASWY TSAASSINSH LPPPGSLQHP LALASGTLV

```

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1                   1.     A method of detecting a lung cancer-associated transcript in a cell  
2     from a patient, the method comprising contacting a biological sample from the patient with a  
3     polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence  
4     as shown in Tables 1A-16.
- 1                   2.     The method of claim 1, wherein the polynucleotide selectively  
2     hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-16.
- 1                   3.     The method of claim 1, wherein the biological sample is a tissue  
2     sample.
- 1                   4.     The method of claim 1, wherein the biological sample comprises  
2     isolated nucleic acids.
- 1                   5.     The method of claim 4, wherein the nucleic acids are mRNA.
- 1                   6.     The method of claim 4, further comprising the step of amplifying  
2     nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1                   7.     The method of claim 1, wherein the polynucleotide comprises a  
2     sequence as shown in Tables 1A-16.
- 1                   8.     The method of claim 1, wherein the polynucleotide is labeled.
- 1                   9.     The method of claim 8, wherein the label is a fluorescent label.
- 1                   10.    The method of claim 1, wherein the polynucleotide is immobilized on  
2     a solid surface.
- 1                   11.    The method of claim 1, wherein the patient is undergoing a therapeutic  
2     regimen to treat lung cancer.
- 1                   12.    The method of claim 1, wherein the patient is suspected of having lung  
2     cancer.
- 1                   13.    A method of monitoring the efficacy of a therapeutic treatment of lung  
2     cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and

5 (ii) determining the level of a lung cancer-associated transcript in the  
6 biological sample by contacting the biological sample with a polynucleotide that selectively  
7 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16,  
8 thereby monitoring the efficacy of the therapy.

1 14. The method of claim 13, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated transcript to a level of the lung cancer-associated  
3 transcript in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 15. The method of claim 13, wherein the patient is a human.

1 16. A method of monitoring the efficacy of a therapeutic treatment of lung  
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and

5 (ii) determining the level of a lung cancer-associated antibody in the biological  
6 sample by contacting the biological sample with a polypeptide encoded by a polynucleotide  
7 that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in  
8 Tables 1A-16, wherein the polypeptide specifically binds to the lung cancer-associated  
9 antibody, thereby monitoring the efficacy of the therapy.

1 17. The method of claim 16, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated antibody to a level of the lung cancer-associated  
3 antibody in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 18. The method of claim 16, wherein the patient is a human.

1 19. A method of monitoring the efficacy of a therapeutic treatment of lung  
2 cancer, the method comprising the steps of:

3 (i) providing a biological sample from a patient undergoing the therapeutic  
4 treatment; and



5 (ii) determining the level of a lung cancer-associated polypeptide in the  
6 biological sample by contacting the biological sample with an antibody, wherein the antibody  
7 specifically binds to a polypeptide encoded by a polynucleotide that selectively hybridizes to  
8 a sequence at least 80% identical to a sequence as shown in Tables 1A-16, thereby  
9 monitoring the efficacy of the therapy.

1 20. The method of claim 19, further comprising the step of: (iii) comparing  
2 the level of the lung cancer-associated polypeptide to a level of the lung cancer-associated  
3 polypeptide in a biological sample from the patient prior to, or earlier in, the therapeutic  
4 treatment.

1 21. The method of claim 19, wherein the patient is a human.

1 22. An isolated nucleic acid molecule consisting of a polynucleotide  
2 sequence as shown in Tables 1A-16.

1 23. The nucleic acid molecule of claim 22, which is labeled.

1 24. The nucleic acid of claim 23, wherein the label is a fluorescent label

1 25. An expression vector comprising the nucleic acid of claim 22.

1 26. A host cell comprising the expression vector of claim 25.

1 27. An isolated polypeptide which is encoded by a nucleic acid molecule  
2 having polynucleotide sequence as shown in Tables 1A-16.

1 28. An antibody that specifically binds a polypeptide of claim 27.

1 29. The antibody of claim 28, further conjugated to an effector component.

1 30. The antibody of claim 29, wherein the effector component is a  
2 fluorescent label.

1 31. The antibody of claim 29, wherein the effector component is a  
2 radioisotope or a cytotoxic chemical.

1 32. The antibody of claim 29, which is an antibody fragment.

- 1                    33.    The antibody of claim 29, which is a humanized antibody
- 1                    34.    A method of detecting a lung cancer cell in a biological sample from a  
2 patient, the method comprising contacting the biological sample with an antibody of claim  
3 28.
- 1                    35.    The method of claim 34, wherein the antibody is further conjugated to  
2 an effector component.
- 1                    36.    The method of claim 35, wherein the effector component is a  
2 fluorescent label.
- 1                    37.    A method of detecting antibodies specific to lung cancer in a patient,  
2 the method comprising contacting a biological sample from the patient with a polypeptide  
3 encoded by a nucleic acid comprises a sequence from Tables 1A-16.
- 1                    38.    A method for identifying a compound that modulates a lung cancer-  
2 associated polypeptide, the method comprising the steps of:  
3                    (i) contacting the compound with a lung cancer-associated polypeptide, the  
4 polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least  
5 80% identical to a sequence as shown in Tables 1A-16; and  
6                    (ii) determining the functional effect of the compound upon the polypeptide.
- 1                    39.    The method of claim 38, wherein the functional effect is a physical  
2 effect.
- 1                    40.    The method of claim 38, wherein the functional effect is a chemical  
2 effect.
- 1                    41.    The method of claim 38, wherein the polypeptide is expressed in a  
2 eukaryotic host cell or cell membrane.
- 1                    42.    The method of claim 38, wherein the functional effect is determined by  
2 measuring ligand binding to the polypeptide.
- 1                    43.    The method of claim 38, wherein the polypeptide is recombinant.

1           44.    A method of inhibiting proliferation of a lung cancer-associated cell to  
2    treat lung cancer in a patient, the method comprising the step of administering to the subject a  
3    therapeutically effective amount of a compound identified using the method of claim 38.

1           45.    The method of claim 44, wherein the compound is an antibody.

1           46.    The method of claim 45, wherein the patient is a human.

1           47.    A drug screening assay comprising the steps of

2           (i) administering a test compound to a mammal having lung cancer or a cell  
3    isolated therefrom;

4           (ii) comparing the level of gene expression of a polynucleotide that selectively  
5    hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-16 in a  
6    treated cell or mammal with the level of gene expression of the polynucleotide in a control  
7    cell or mammal, wherein a test compound that modulates the level of expression of the  
8    polynucleotide is a candidate for the treatment of lung cancer.

1           48.    The assay of claim 47, wherein the control is a mammal with lung  
2    cancer or a cell therefrom that has not been treated with the test compound.

1           49.    The assay of claim 47, wherein the control is a normal cell or mammal.

1           50.    A method for treating a mammal having lung cancer comprising  
2    administering a compound identified by the assay of claim 47.

1           51.    A pharmaceutical composition for treating a mammal having lung  
2    cancer, the composition comprising a compound identified by the assay of claim 47 and a  
3    physiologically acceptable excipient.

